Occupational Choice under Credit and Information Constraints*

Adam Osman †

Yale University

January 16, 2014

Abstract

Credit and information constraints can affect not only participation levels in different occupations, but also the types of individuals found in those occupations. I develop a model of occupational choice which shows how new information alters expectations and thus occupational choice. Credit and information constraints also interact with each other: the *same* information can have opposing impacts on occupational choice depending on the presence of credit constraints. Using a survey and information experiment in seven vocational high schools in Egypt, I find support for the model's predictions, including a key compositional prediction: risk averse individuals respond more to changes in expectations of risk at both the intensive (variance of income) and extensive (probability of finding a job) margins. This differential response leads highly risk averse individuals to shift towards "safer" occupations on average, and out of using credit to start a small enterprise. Because those who are risk averse report lower returns to credit in general (by favoring lower risk/lower return investments), the average expected income for those utilizing credit rises by 12% relative to the no information case. I discuss other areas where information interventions may influence participation in, and thus the efficacy of, complementary policies.

Keywords: Occupational Choice, Subjective Expectations, Information, Credit Constraints, Entrepreneurship

JEL Codes: J24, L26, E24, G14, O12, O15

*The author would like to thank Dean Karlan, David Atkin, Costas Meghir, as well as Chris Udry, Mark Rosenzweig, Amit Khandelwal, Basit Zafar, Jamin Speer, Muneeza Alam, Camilo Domingez, Treb Allen and Kareem Haggag. The author would also like to thank the Misr El Kheir Foundation, the Drosos Foundation, and the Economic Research Forum, Abdel Rahman Nagy, Mahmoud Saeed, Fyras Mawazini, Ragui Assaad, Hala Aboulsayed, Shereen Kandil and Muna Zughayer. This research was supported by a Sasakawa Research Grant and the Kaufmann Dissertation Fellowship. Approval was obtained from the Yale University Human Subjects Committee, IRB1211011109.

†Email:Adam.Osman@Yale.edu; The most recent version of this paper can be found at http://www.adam-osman.com

1 Introduction

The study of occupational choice has a long history dating back at least to the Roy model and continues to be an active area of research in a variety of subfields in economics.[1] Occupational choice has been shown to impact issues including wage inequality, economic growth, and employment, leading to interest in the subject from policy makers as well as researchers.[2] More recent work has shown the importance of credit constraints on the occupational decisions of individuals, an issue especially important in developing economies, but none have considered the importance of information constraints.[3] Yet, information constraints have been shown to have significant impacts on decision making.[4]

I develop a model of occupational choice that accounts for both credit constraints and information constraints. The model allows for an individual's expectations to be informed by their personal information set, which may or may not be accurate.[5] It also helps explain the important ways in which credit and information constraints can interact– in particular, the *same* information can have opposing effects on occupational choices depending on the presence of credit constraints. I test the model using a unique survey and information experiment in seven vocational high schools in Egypt. The data I collect support the model's predictions, including a key compositional prediction: risk averse individuals respond more strongly to changes in expectations about the variance of income in an occupation and the probability of succeeding in the occupation (i.e. actually finding a job conditional on looking). This differential response leads highly risk averse individuals to shift towards "safer" occupations on average, and out of using credit to start a small enterprise. Because those who are risk averse report lower returns to credit in general (by favoring lower risk/lower return investments), the average expected income for those utilizing credit rises by 12% relative to the no information case.

More specifically, the model outlines how new information can impact an individual's expectations, and in turn, change their occupational choices. There is no unified use of the term occupation in the literature and so in my model I define three: (1) Wage Work, (2) Entrepreneurship, and (3) Inactivity. The model shows that information which leads to changes in expectations which increase (decrease) expected utility in one occupation relative to another, leads to increases (decreases) in the intent to pursue that occupation relative to the other. Changes in expectations about risk at the intensive (variance of income), and extensive (probability of finding a job) margins, also lead

[1] See e.g. Roy (1951); Ginzberg et al. (1951); Blau et al. (1955)

[2] See e.g. Heckman et al. (1998); Banerjee and Newman (1993);WDR: World BankStaff (2013)

[3] See e.g. Boskin (1974); Evans and Jovanovic (1989); Banerjee and Newman (1993); Bianchi and Bobba (2013)

[4] See e.g. Duflo and Saez (2003); Jensen (2010); Nguyen (2008); Bettinger et al. (2012)

[5] Although individuals will most certainly have information about themselves that is accurate, and not available to the econometrician, they may also hold beliefs about known values that are incorrect- for instance the population average income in an occupation.

to corresponding changes in occupational intentions. Using a utility function that accounts for risk aversion leads to the prediction that highly risk averse individuals respond more strongly to changes in expected risk than those who are not highly risk averse.

The model also helps make clear the identification issues inherent in the estimation of the determinants of occupational choice. In particular, the model outlines the biases that may arise if unobservable taste parameters are correlated with both the occupational expectations and occupational intentions. I use the information intervention to create a panel of expectations and occupational intentions, similar to Wiswall and Zafar (2013). This panel allows me to difference out any time-invariant taste parameters that may otherwise bias my results, improving upon earlier strategies for estimating the determinants of occupational choice.

To test the model I conduct a survey and information experiment with seniors in seven vocational high schools in Al-Fayoum, Egypt, a city two hours from Cairo. Many of the students are credit constrained (coming from generally poor families) and information constrained (due to a lack of access to official labor market statistics), making this context well suited for testing the model. These students almost never go on to pursue higher education, due to the structure of the Egyptian education system. Hence, the sample is soon to enter the labor market, making the occupational choice decision a very relevant one.

The survey and experiment is comprised of five main stages. The first stage collects basic demographic information about the students, as well as measures of risk aversion. The second stage elicits the students' expectations of outcomes in the different occupations (wage work, entrepreneurship and inactivity). In particular, it focuses on beliefs about average income in the population in an occupation and then moves to expectations of personal income, variance of income, and the probability of success.[6] To measure the variance of income, the survey elicits from the student's the probability density function of their expected earnings. It does this using five "bins" of different income levels and asks the students to allocate 100 "points" across the bins to reflect what they believe is the probability they would receive that level of income, conditional on successfully pursuing that occupation. After the elicitation of expected outcomes in the different occupations, students report their occupational intentions- the probability that they will pursue each occupation when they enter the labor market. The survey then introduces a *hypothetical* credit intervention, which mirrors a credit intervention that is being planned by a local NGO, that provides enterprise credit to those who want to start a business.[7] Students are asked to provide their expectations of outcomes in entrepreneurship while having access to this loan, as well as their

[6]"Probability of success" refers to the probability of finding a job conditional on pursuing wage work/probability of starting a successful business conditional on pursuing entrepreneurship.

[7]The credit intervention is a loan of 10,000LE (approximately $1450) that must be paid back over the course of a year. The interest rate on the loan is zero to avoid any issues of individuals refusing the loan on religious grounds, due to the Islamic impermissibility of dealing with interest (El-Gamal et al., 2013).

occupational intentions given the availability of credit.

The third stage of the survey occurs directly after the elicitation of expectations of outcomes and occupational intentions, with and without credit. The students are provided information about the distribution of incomes in the different occupations. These data come the Egypt Labor Market Panel Survey (ELMPS2012), the first representative survey of the Egyptian population post-revolution. Unlike earlier research on information interventions, which simply provided individuals with averages or estimated returns, I provide the full distribution of income to the students. This allows for an arguably more accurate representation of labor market outcomes, and a greater emphasis on the natural heterogeneity of income. The students are also provided with the proportion of the working population in each of the occupations and a short explanation of the endogeniety of the income[8]. The data I provide are restricted to those who have the same level of education as the students and are in the same general age range (18-30).

The fourth stage of the survey occurs after the information intervention. Students are asked again to provide their expectations of outcomes in the different occupations and the corresponding updates to their occupational intentions. This allows for the measurement of how the expectations and intentions were impacted by the information intervention. The credit intervention is then re-introduced to measure how the response to credit has changed after the information intervention. This leads to four main cells of occupational expectations and intentions: (1) Baseline, (2) with Credit, (3) with Information, (4) with Credit *and* Information. The fifth stage of data collection is a survey of the students after they graduate, allowing for the measurement of how their actions in the labor market line up with their reports in the first four parts of the survey. Finally, a group of students were surveyed but not provided any information. Following up with these students after they graduate, and comparing their outcomes with those that received the information, allows for an explicit test of the reported effects of the information intervention. The students that did not receive the information serve as an important comparison group and robustness check.[9]

Moving to analysis of the data, I find that the information intervention successfully imparted information about incomes in the occupations to the students, leading to changes in beliefs about the average income in the population and, in turn, changes in expectations of personal income. I use these changes in expectations to test the model's implications. As predicted, when the infor-

[8]The exact explanation of the endogeneity of income can be found in the appendix below. It attempts to make clear that the data are the result of choices by individuals about their most preferred option. The explanation also included the concept that the best self-employed person might not be the best wage worker if they changed occupations and vice versa.

[9]The survey was implemented at the classroom level. The allocation of classes to the information/no-information group was not explicitly random. The classes that received the information were chosen on an availability basis (the classrooms that were free during the time of the survey) and the classes that did not receive the information were returned to later. Although less than ideal, this quasi-random selection leads to groups that are equal on observables. The data from the no-information group are used primarily as an important robustness check for the changes in actions reported by the sample due to the information.

mation leads to increases (decreases) in expected utility in one occupation, relative to the other occupations, individuals increase (decrease) their intent to pursue that occupation. The data also confirm that highly risk averse individuals respond much more strongly to changes in expectations of risk in an occupation at both the intensive (variance of income), and extensive (probability of finding a job) margins.

In the schools, the information intervention leads to an aggregate shift of the sample away from inactivity and towards entrepreneurship. The credit intervention leads to a shift into entrepreneurship larger than the one induced by the information intervention. Information and credit together also lead to a shift into entrepreneurship but a *smaller* shift than just the credit intervention. That is, although information alone leads to a shift into entrepreneurship, the *same* information, when credit is available, leads to a shift *away* from entrepreneurship. I explain these aggregate shifts by combining my estimates of the determinants of occupational choice with the reported changes in expectations by the sample. I find that the information intervention has different effects on expectations in the credit constrained vs. unconstrained case, explaining why it can have opposing aggregate impacts on occupational intentions. A clarifying example is to consider an individual who believes that their average income in entrepreneurship when credit constrained is 500LE and 2500LE when credit is available. If they receive information that shows that all those in entrepreneurship make between 1000 and 2000LE a month, it would lead to an increase in expectations of income in the constrained case and a decrease in the unconstrained case. Follow up data show that the reported changes from the information intervention are mirrored in the actions taken by the sample when they enter the labor market, lending credibility to the changes reported in the survey.

The aggregate effects mask important compositional changes that occur due to the differential reaction of highly risk averse individuals to changes in expectations about risk. While the information intervention leads to an overall shift into entrepreneurship, those who are highly risk averse instead shift into wage work. Similarly, when both credit and information constraints are relaxed, the highly risk averse shift away from entrepreneurship more than the rest of the sample. Evidence also shows that the highly risk averse plan to use the credit made available to them on lower risk/lower return activities.

From a policy standpoint, the information intervention leads to a more efficient allocation of individuals across occupations, by inducing individuals to change their occupational intentions. This improves individual welfare on average by 6.5%. The relaxation of credit constraints increases expected income and improves sector allocation leading to a 10% increase in welfare. Relaxing both constraints leads to a 12.3% increase in welfare through both increased expected income and improvements in occupational intentions. These changes in welfare are even larger when using a utility function that accounts for risk aversion. The information intervention also leads to changes in the utilization of the hypothetical credit intervention. The proportion of the sample that reports that they would use the credit intervention drops from 30.0% before the information intervention to 24.2% after the information intervention. The proportion of highly risk averse individuals utilizing credit also drops from 30.3% to 23.1%. Because highly risk averse individuals report lower returns to credit (by favoring lower risk/lower return activities), the average income in the credit intervention *increases* by 11.5% relative to the no information case.

These results have significant implications for policy makers as well as researchers. First, information can be sufficient to change occupational choices. These changes can be explained by changes in expectations that are induced by the new information. Second, highly risk averse individuals respond more strongly to changes in expectations of the variance of income and the probability of success in an occupation. This differential reaction leads to a change in the composition of individuals across occupations, further improving individual welfare. Third, information and credit constraints interact in a way that impacts the efficacy of the credit intervention, leading to improved selection into its utilization and increasing the average expected income of those that would use it.

Earlier work on information interventions rarely consider how the information interacts with the constraints facing the individuals (Duflo and Saez, 2003; Jensen, 2010; Nguyen, 2008; Dupas, 2011). This paper shows that these interactions can be non-trivial and have important impacts on the outcomes of the individuals as well as the outcomes of the policy tool being studied. Other work considers the importance of getting credit to *transformational* entrepreneurs who have the highest return to credit and who start businesses that generate employment (De Mel et al., 2011). In this vein, my results show how a careful bundling of information can allow for low-return individuals to self-select themselves out of utilizing credit and improve the efficacy of a credit intervention.

Governments and international organizations commonly promote entrepreneurship as a solution to unemployment. For instance, Ban Ki-Moon, Secretary General of the United Nations, recently said, "Entrepreneurship can be part of the solution by transforming unemployed young people into major employers" Ki-Moon (2013). Policy makers of this opinion want to encourage individuals to start businesses, but to do so successfully, they must first understand the determinants of occupational choice. My results shed light on the determinants of occupational choice as well as how the relaxation of credit and information constraints can impact both the proportion and composition of individuals in entrepreneurship.

The paper proceeds as follows: Section 2 reviews the relevant prior literature, Section 3 puts forth a model of occupational choice and derives its testable implications, Section 4 describes the details of the survey experiment, the elicitation procedures and provides details about the sample and local context, Section 5 presents the results of the paper while Section 6 concludes with an overview, policy implications, and directions for future work.

2 Prior Literature

2.1 Occupational Choice

Occupational choice impacts a variety of issues ranging from the wage inequality to issues of sector allocation, personnel economics and economic growth (Roy, 1951; Ginzberg et al., 1951; Blau et al., 1955; Miller, 1984; Doepke and Zilibotti, 2008). Occupational choice's importance in economic development hasn't been ignored, with many papers focusing on the relationship between economic development and entrepreneurship (Banerjee and Newman, 1993; Ghatak and Nien-Huei Jiang, 2002; Carree and Thurik, 2005; Naudé, 2010; Eeckhout and Jovanovic, 2012). Many governments and international organizations such as the World Bank and the IMF sometimes promote entrepreneurship as a "solution to unemployment" (Staff, 2013). These organizations often advocate increasing the proportion of the population who should pursue entrepreneurship[10] but to do so successfully we must first understand the main determinants of this choice of occupation.

Credit constraints can impede the ability of individuals to pursue entrepreneurship even if they desire to do so as discussed in (Banerjee, 2001; Bianchi and Bobba, 2013; Evans and Jovanovic, 1989; McKenzie and Woodruff, 2008; Blattman et al., 2013). Recent work has shown that credit constraints can have important effects on business outcomes (Banerjee et al., 2013; Crépon et al., 2011; Karlan and Zinman, 2010, 2011; Augsburg et al., 2012; Attanasio et al., 2011; Angelucci et al., 2013). It is natural that constraints of this type will impact the outcomes an individual will expect to have if they choose to pursue entrepreneurship.

The relationship between risk aversion and occupational choice has been written about for decades, (King, 1974; Siow, 1984). It is intuitive that if different occupations come with different levels of risk then an individual's appetite for risk will serve as an important determinant of which occupation they will pursue. There are also efficiency concerns for the economy at large when considering risk aversion. We might think that there is ideal allocation of risk-taking across sectors that policy makers might strive to achieve. This allocation would intuitively require less risk-averse individuals to pursue the occupations that require risk taking (Jeong and Townsend, 2007).

It has also been shown that information can affect the occupational choice decisions of individuals. Papers like De Mel et al. (2011, 2012); Klapper and Love (2011) show that information and information interventions can impact the way people approach their choice of occupation and eventually their business decisions.

[10] Whether or not promoting entrepreneurship would in fact lead to improved economic growth is outside the scope of this paper. Nonetheless understanding what might induce an individual to pursue or leave entrepreneurship can still be informative.

2.2 Information and Subjective Expectations

New information has been shown to have important effects on decision making. Papers like Jensen (2010); Duflo and Saez (2003); Nguyen (2008); Dupas (2011); Bettinger et al. (2012)show relatively large effects of information provision on topics ranging from education to HIV testing to retirement savings.

It is natural to tie the literature on the impacts of new information to the literature on subjective expectations. Although directly eliciting expectations from individuals had been seen as problematic for some time, recent work has shown that subjective expectation data can be informative (Manski, 2004; Dominitz and Manski, 1997). Subjective expectations have been shown to be useful in studying a variety of topics including health and education (Attanasio and Kaufmann, 2009; Delavande, 2008a; Delavande and Kohler, 2009; McKenzie et al., 2013; Stinebrickner and Stinebrickner, 2013; Kaufmann, 2009).

With particular relevance to this paper Zafar (2011a) , Arcidiacono et al. (2012) and Wiswall and Zafar (2013) show that eliciting expectations can inform us about models of discrete choice. Their work considers the college major decision and is as close as we get to how expectations factor into occupational choice [11] .Wiswall and Zafar (2013) are especially notable in introducing a technique to deal with the biases that come from being unable to control for unobservable tastes parameters in a model of choice of college major that utilizes subjective expectations. They show how collecting expectations and choices before and after an information intervention can allow for the creation of a panel of beliefs which can be used to difference out any time-invariant unobservables that may otherwise bias estimation. Arcidiacono et al. (2013) use the data collected from their earlier paper (Arcidiacono et al., 2012) to consider the *ex-ante* expected treatment effects of choosing different occupation types in a sample of male undergraduates at Duke.

3 A Model of Occupational Choice

3.1 Set Up

In this model individuals have to choose between pursuing one of three occupations. Labor is indivisible and so individuals cannot choose more than one occupation. I define the three occupations as: (1) Wage Work, (2) Entrepreneurship (self-employment) or (3) Inactivity (leaving the labor force). In the simplest case an individual compares their expected utility in each of the three occupations and chooses the one with the highest value. Hence the optimization problem can be written as:

[11] Keats (2012) considers the *secondary* occupational choice decisions of agricultural workers in Kenya but does not manage to deal with the endogeniety biases found in working with expectations data.

$$\max_{d=\{w,s,l\}} E[U_d = u(X_d, \gamma_d, \rho)|\Omega_t]$$

Each occupation has a corresponding utility that depends on three main factors- (1) a vector of observable inputs X, which includes things like income, personal and family characteristics, etc, (2) a vector of unobservable inputs γ, which includes things like tastes for non-pecuniary aspects of occupations, and (3) behavioral preferences, in particular risk aversion, ρ. Unlike in earlier models of occupational choice I explicitly account for the individual's information set, Ω_t, as I allow this to change over time and impact the individual's expectations.

In this setting it is possible for an individual to be "information constrained". This individual can be missing some information that may change their expectations of future outcomes if they had access to it. In earlier models of occupational choice the only way to shift occupational intentions was to change the underlying returns of an activity. This model allows for information alone to induces changes in actions. In section 3.4 below I derive exactly how I expect new information to affect occupational choice through changes in expectations.

Income is an important observable input. Let it be defined as $I_d = f(\theta, \lambda, C)$, where θ represents the individual's ability (High or Low), λ represents the individual's credit constraint, and C represents the individual's other personal characteristics. Income will take slightly different forms for each of the three occupations. In the most general way we have:

$$I_d = \begin{cases} g(\theta, C) & \text{if Wage Work} \\ h(\theta, \lambda, C) & \text{if Self-Employment} \\ j(C) & \text{if Left Labor force} \end{cases}$$

This is simply to say that credit constraints only matter when considering income in entrepreneurship and ability does not matter when considering the effective income in leaving the labor force. Personal characteristics, including basic demographics, can affect income in all of the occupations.

Finally, risk preferences enter the model through the specific form of the utility function. In an effort to keep the utility function general I can simply assume that risk aversion leads to lower utility for incomes that are riskier on the intensive (more volatile) and extensive (larger probability of failing in the occupation) margins. A variety of functional forms generate this relationship including the most common utility functions that account for risk aversion[12].

[12]E.g. CRRA, Log Utility, etc.

3.2 Model Timing

At all points in time individuals do not know with certainty what their income will be in any of the potential occupations. Instead, they have a set of subjective expectations about what their income might be. Unlike in earlier models of occupational choice I make no assumptions about the rationality of these expectations. Expectations in this model depend on the individual's current information set and the way in which they transform that information into their expectations.[13] As the information set changes the expectations an individual holds can also change.

There are three time periods in the model, t=0, when the individual is considering which occupation to pursue utilizing their baseline beliefs and expectations; t=1, when the individual considers which occupation to seek given their beliefs after they've received new information; and, t=2 when the individual has to decide which occupation they will be seeking as they enter the labor market. In the empirical section below the difference between t=0 and t=1 will be a couple of hours whereas the difference between t=1 and t=2 will be several months.

At t=0 & t=1, there is still time before the individual has to make their final decision. In that time there is the potential for some uncertainty (referred to as resolvable uncertainty in Blass et al. (2010)) that leads to the optimization problem being solved by a set of probabilities of choosing an occupation at t=2, as opposed to a simple discrete choice between occupations. This type of uncertainty is intuitive - it is possible that between the time when the individual is considering which occupation to pursue and when they are forced to choose an occupation many things can change. Changes can range from issues like local economic conditions, changes in family circumstances, or changes in individual occupational tastes. At t=2 the resolvable uncertainty is realized and the individual chooses the occupation that has the highest expected utility and pursues it (labor is indivisible and so individuals can only choose to pursue one particular occupation at a time).

3.3 Solution to the Model

In most of the empirical work to follow I will be looking at the probability that the individuals have, at t=0 & t=1, regarding their eventual occupational choice decision. Following the well-known solution to these types of discrete choice models if we assume i.i.d. extreme value errors (which is without loss of generality as per McFadden and Train (2000)) then the probability of

[13]Expectations do not depend on information alone. Two individuals with the same information might have different expectations if the function that transforms their information set into their expectations differ. Since I collect data on the individual's expectations directly I need not assume or model how a change in information translates into a change in expectations. Instead, I can simply measure the changes in expectations directly and consider how those changes impact occupational choice.

choosing occupation d at time t=0 or t=1 is Y_d:

$$Y_{dt} = \frac{exp\{E[U_d|\Omega_t]\}}{\sum_{dt} exp\{E[U_d|\Omega_t]\}} \tag{1}$$

Whereas when t=2, the solution to the model will more simply entail the contrasting of the expected utilities in each of the occupations and choosing the one with the highest value.

Throughout this paper, I am interested in the occupations that individuals plan to initially *pursue* when it is their time to enter the labor market (in this context right after they graduate from high school). It is important to note that one's intention to seek wage work, for instance, is not the same as finding a wage job. Nonetheless it is reasonable to assume that an increase in the probability of pursuing wage work, ceretis parabis, will result in an increase in the probability of ending up in a wage job. The empirical work below will provide evidence that this is in fact the case. This allows us to treat changes in the intention to pursue a particular occupation as similar to changes to the probability of being in that occupation at some future point in time.

3.4 Implications of Relaxing Information Constraints

The effects of providing information naturally depend on both the content of the information as well as the initial information set of the individual. In the extreme case where the information does not provide any new knowledge, I would expect to find no effects at all. Intuitively, providing information that is redundant should lead to no changes in expectations nor in the behavior of individuals. If, on the other hand, the information was novel and managed to change the individual's expectations about future outcomes, then it might change an individual's occupational intentions.

In the model I make no assumptions about how new information will lead to changes in expectations. It is sufficient that new information can change expectations in some way. I then work directly with the expectations themselves and consider how changes in expectations lead to changes in occupational choices. This allows me to forgo making any assumptions about how individuals update their expectations when they encounter new information.

If I consider the simplest case, it is clear that all else equal, an increase in the expected income in, for instance entrepreneurship, should lead to an increase in the probability of pursuing entrepreneurship (as can be seen from equation [1]). On the other hand, if the increase in entrepreneurship is relatively less than a corresponding increase in the expected income in wage work, then I would expect the probability of pursuing entrepreneurship to decrease. It is not sufficient to consider changes in expectations in one occupation without considering the changes in expectations of other occupations.

More formally I can find that relative to a reference occupation, an increase in expected utility in one occupation will lead to an increase in the probability of pursuing that occupation. Following Wiswall and Zafar (2013), this can be best shown by taking logs and comparing two different occupations before and after an information intervention.

First consider the case without the information intervention. Let $r_{dt,d't}$ be the log ratio of the intention to pursue occupation d relative to occupation d' at time t.

$$r_{dt,d't} = log(Y_{dt}) - log(Y_{d't})$$

This allows the simplification of the expression from equation [1] above to

$$= \beta(E[U_d|\Omega_t] - E[U_{d'}|\Omega_t])$$

If I assume X and γ are additively separable, then

$$= \beta_1(E[u(X_d|\Omega_t)] - E[u(X_{d'}|\Omega_t)]) + \beta_2(E[u(\gamma_d|\Omega_t)] - E[u(\gamma_{d'}|\Omega_t)]) \tag{2}$$

This relationship can be estimated if the data is available. Unfortunately the γ's are unobservable and if they are correlated with both Y and X then they will lead to omitted variable bias when attempting to estimate the β's.[14]

If I assume that the unobservables are time-invariant, then by introducing new information, and collecting data before and after the intervention, I will be able to difference out the unobservables and estimate this relationship without the omitted variable bias. This assumption is not unreasonable if the difference in time between the first observation and the second is small and the new information informs only expectations of observables. In the empirical work below the difference in time between t=0 & t=1 will only be a couple of hours and the information I provide is completely about observable characteristics of occupations.

More formally I can look at the difference in $r_{dt,d't}$, the log ratio of occupation d to occupation d' over time:

$$\Delta r_{d,d'} = r_{dt',d't'} - r_{dt,d't} = (lnY_{dt'} - lnY_{d't'}) - (lnY_{dt} - lnY_{d't})$$
$$= \beta_1[(E[u(X_d|\Omega_{t'})] - E[u(X_{d'}|\Omega_{t'})]) - (E[u(X_d|\Omega_t)] - E[u(X_{d'}|\Omega_t)])] \tag{3}$$

With the assumption of time-invariant unobservables I difference them out when contrasting $r_{d,d'}$at time t and $r_{d,d'}$at time t'. This provides equation [3] which is now able to be estimated without omitted variable bias from the unobservables.

With this I can now show that

[14]Examples of the type of correlations and bias can be found in section 5.2 below.

$$\frac{\partial r_{d,d'}}{\partial(E[I_d]-E[I_{d'}])} > 0$$

That is, the model predicts that as the income in occupation d *relative* to the income in occupation d' increases then the intent to pursue occupation d *relative* to the intent to pursue occupation d'will increase (assuming that utility is increasing in income). Similarly if I assume that utility weakly decreases with the variance of income then I can also conclude that as the variance of income in occupation d, *relative* to the variance of income in occupation d' increases then the intent to pursue occupation d *relative* to the intent to pursue occupation d'will weakly *decrease:*

$$\frac{\partial r_{d,d'}}{\partial(E[Var(I_d)]-E[Var(I_{d'})])} \leq 0$$

Intuitively I would expect some of these effects to be more pronounced for individuals who are risk averse. For instance as the variance of income increases in one occupation relative to another I expect that the intent to pursue that occupation will decrease but I also expect that it will decrease *more* for individuals who are particularly risk averse. To show this I need to assume a functional form for the utility function that allows for risk aversion. Taking the standard CRRA form would lead to the following update to equation [3]:

$$= \beta_1[(E[\frac{X_{dt'}^{1-\rho}}{1-\rho}]-E[\frac{X_{d't'}^{1-\rho}}{1-\rho}])-(E[\frac{X_{dt}^{1-\rho}}{1-\rho}]-E[\frac{X_{d't}^{1-\rho}}{1-\rho}])] \tag{4}$$

In the empirical work below I will consider the changes in the log of the observable inputs like income and variance. This will allow for easier interpretation of the coefficients as elasticities of changes in expectations into changes in intentions for occupations. In this case, by taking logs of the utility and using its properties I would then have:

$$= \beta_1(1-\rho)\{(E[\log X_{dt'}|\Omega_{t'}]-E[\log X_{d't'}|\Omega_{t'}])-(E[\log X_{dt}|\Omega_t]-E[\log X_{d't}|\Omega_t])\} \tag{5}$$

In the empirics below all individuals will be considered risk averse whereas a proportion will be considered *highly* risk averse. As expected, equation [5] indicates that the effects of changes in expectations on occupational choice can differ based on one's level of risk aversion.

To summarize: in cases in which information induces a change in expected utility those changes will result in logical revisions in intentions to pursue a particular occupation relative to the changes in another occupation. To take a concrete example from the empirics to follow- as income in wage work increases relative to income in entrepreneurship I expect to see an increase in the intention to pursue wage work relative to the intention to purse entrepreneurship. For inputs that cause utility in an occupation to decrease I expect to find an increase in that input (like the variance of income) will result in a decrease in the intent to pursue that occupation relative to another. Finally, I expect that the sensitivity to changes in expectations differ for those with different levels of risk aversion.

3.5 Implications of Relaxing Credit Constraints

The implications of relaxing credit constraints are more straight forward. It's important to remember that the form of the income function in section 3.1 had credit constraints enter the model only through income in entrepreneurship. A needed, yet nonrestrictive, assumption is that income in entrepreneurship weakly increases as credit constraints are relaxed. A standard way to show this is to have income in entrepreneurship depend on the optimal capital allocation of the individual. This can be determined by solving a simple optimization problem of the following form:

$$\max_k p * f(k, \theta) - (k - W)r$$

$$\text{s.t. } 0 \leq k \leq \lambda W$$

Where p is the price of the good, f is the production function, and W is starting wealth. The constraint is that k can only be as large as λW where $\lambda \geq 1$ and depends on the credit constraints the individual faces. In the most constrained case $\lambda = 1$ and the individual can't access any additional financing.

Given this set up, it is easy to see that as I relax credit constraints I will weakly increase the income in entrepreneurship for the sample and nothing else. This should then increase the probability of an individual pursuing entrepreneurship relative to any of the other occupations.

Hence the model predicts that

$$\frac{\partial U_s}{\partial \lambda} \geq 0 \text{ and hence } \frac{\partial Y_s}{\partial \lambda} \geq 0$$

Individuals in the sample should be more likely to intend to pursue entrepreneurship when given the option of additional credit access if they are credit constrained.

3.6 The interaction of information and credit constraints

The interaction of information and credit constraints can have important consequences. As an example, it is not inconsistent for the same information intervention to lead to an increase in an individual's intention to pursue self employment when credit is not available while leading to a decrease in their intent to pursue self employment when credit is available.

The model in its current state is capable of explaining such an event without having to directly model the interaction of the two constraints. This is because, as section 3.4 made clear, the impact new information has on occupational intentions comes through its impact on the expectations of utility in the occupations. The earlier example could be rationalized simply by finding that the information changed expectations differently when credit was available as opposed to when it was not. The model does not make any assumptions about how individuals use the new information to update their expectations and so this becomes a purely empirical phenomenon.

A concrete example that will make this clear considers utility in self employment. An individual must consider income when credit constrained ($I_{\lambda=1}$) and income when unconstrained ($I_{\lambda>1}$). Information can act on these expected incomes differently. Imagine the case where at baseline an individual has low expectations of income in entrepreneurship when credit constrained ($I_{\lambda=1}=500$) and high expectations of income in entrepreneurship when not credit constrained (say $I_{\lambda>1}=2500$). That individual might be provided with information about the distribution of income in the general population that leads them to revise their expectations of income in self employment when credit constrained upwards while revising their expectations of income in self employment when unconstrained downwards ($I_{SE}\varepsilon[1000,2000]$). Keeping all other expectations equal this would lead to the same information inducing an increase in the intention to pursue self employment in the credit constrained case but a decrease in the intent to pursue self employment in the unconstrained case

4 Data Collection and Experiment Details

This study's sample consists of high school seniors in vocational schools in Al-Fayoum, Egypt. This sample provides a number of advantages for studying these issues of occupational choice.

When considering occupational choice, it is simplest to consider individuals who are about to encounter their first opportunity to enter the labor force. This allows one to effectively ignore any issues of path-dependence or excessive private information about ability or tastes after having tried different potential occupations and succeeding or failing. By studying students near graduation, I manage to avoid these pitfalls and essentially consider occupational choice decision-making with a "clean slate". I can also ignore any issues of differential human capital accumulation as I consider only students with the same level of education. Institutional factors in the Egyptian educational system make it very difficult for vocational high school students to proceed onto higher education, making this occupational choice decision especially important for these students during this time of their lives.

I also expect this sample to potentially be suffering from significant information constraints. This is due to reasons that we would expect students in most low-income countries to be dealing with such as little guidance in school, little access to information about labor market returns and opportunities, and a social stigma associated with discussing incomes. In addition there is an unusually high amount of uncertainty about labor market outcomes for this sample during the study's time period due to the relatively recent government overthrow in early 2011[15]. The uncertainty continued in between the time the students had completed the survey and when they were contacted about their labor market outcomes as a second government overthrow occurred in July 2013.

At the same time, it is not necessary for the students in my sample to be *misinformed* for my identification strategy to be successful. All that is needed is for the sample, at baseline, to not have had access to the data that I provide to them.[16] I describe the data I provide in more detail below but the information was delivered to the students in March/April 2013 while the data set I use was not made public to individuals or researchers until November 2013.

This sample also contains a large number of students who are credit constrained. Al-Fayoum is one of the poorer regions in Egypt and, as will be shown below, the students generally come from poor families. This allows for the study of how credit constraints can effect occupational choice.

Finally by looking at young people I am able to study a sub-population that is of special interest to policy makers. Much has been made of the youth unemployment problem in Egypt, with articles connecting this type of unemployment with the aforementioned government overthrows (Malik and Awadallah, 2013). High youth unemployment is not a problem particular to Egypt. Many countries ranging from others in the Middle East like Jordan and Syria, to Western economies like Spain, France, Greece and the United States suffer from high youth unemployment. A number of governments and international organizations recommend promoting entrepreneurship as a potential solution to this youth unemployment problem making understanding the determinants of occupational choice especially important for this age group.

4.1 Survey Structure

The data was collected over the course of two months in 7 different vocational high schools in Al-Fayoum. The data consists of a three-part survey that was implemented on 1,577 students in their last year of high school.

The first part of the survey covers basic demographic information about the student and their

[15] Anecdotally, during pilot testing, when students were informed that I would be providing them with information about labor market outcomes they would quickly inquire if the data were coming from surveys done before or after the revolution.

[16] In fact even if some of the students had access to the data I provide it would not affect the identification strategy. In that case I should see no change in expectations and similarly no change in occupational intentions for those who already had access to this information. As long as some positive fraction of the students had not seen this data, and it induced them to change their expectations of outcomes, the empirical analysis should still work.

family. It includes information about the student's age, gender, subject of study, and work status. It also includes family information like parental education, parental work status, and information about siblings. Finally, it also includes two questions about risk aversion that I use to characterize those students who are highly risk averse.

The second part of the survey elicits the student's subjective expectations of outcomes in the different occupations after they graduate. It begins by asking the students about what they believe the *population* average monthly income is in wage work for individuals similar to them (recent graduates of 3-year vocational high schools). It then follows up by asking them what they believe their *personal* average monthly earnings would be if they successfully found a wage job after graduation. It goes on to elicit the their beliefs about the distribution of their potential income in wage work post graduation (I explain the elicitation method more thoroughly below). It also inquires about the probability they'd be successful in landing a job if they pursued wage work after they complete their education[17].

The survey then asks the same questions for the case in which they pursue entrepreneurship after graduation. I elicit population and personal averages as well as the distribution of their beliefs about their potential income in entrepreneurship. I again ask what they believe the probability of being able to start a successful business would be after they complete their education. I then ask about their expected income if they leave the labor force after graduation.

After eliciting their expectations of outcomes in each of the occupations the students were then asked to report the probability that they would pursue each of the occupations upon completion of their education. They were supposed to make sure that all of the probabilities in the three occupations added up to 100. For those that did not add up to 100 their responses were normalized to provide effective probabilities of pursuing each occupation.

Finally students were then told to envision having the opportunity to take out a 10,000LE loan which was to be paid back over the course of a year and was conditional on starting a business[18]. They were first asked if they would be interested in taking out such a loan. If so, they were then asked to report how they believe this loan would change their expected income in entrepreneurship, as well as their beliefs about the distribution of income. Finally they were asked once again what their intentions to pursue each occupation would be but this time keeping in mind that they have

[17] We tell the students to provide this information about their expected outcomes after they complete their education. Although very rare some students might want to go on to further school and so might find these questions irrelevant as they are interested in gaining more schooling before they enter the labor market. By asking about their expectations and intentions after they complete their schooling in general (and not just high school) I bypass this problem.

[18] This amount was chosen as it was both very close to the average amount of start up funds used in businesses found in the ELMPS and was very close to the size of the loans that are going to be available to a subset of students in the future as part of a larger experiment on improving youth unemployment. The hypothetical loan is interest free as paying of interest is forbidden in Islam and has led to an under-utilization of conventional credit by Muslims El-Gamal et al. (2013).

access to the loan.

After the second part of the survey, a part of the sample was then provided information about the wage distributions found in the labor market for those with similar characteristics to themselves (in particular those with the same education level and type and were within the ages of 18-30). The details of this information provision can be found below. After the information intervention, the third part of the survey was administered and was identical to the second part. The third part of the survey was meant to measure any differences in expectations and occupational choice intentions that may have occurred due to the new information.

Since the data that is collected is about intended actions, it is not immediately clear if it is as useful as data that considers actual actions. If one can believe the responses of the students about their intentions, then the results I find below will be informative. If on the other hand one does not believe their responses, then the results will have no meaning. For this reason, a set of students were subjected to only the first two parts of the survey and were not given the information treatment. These students serve as a relevant comparison group[19] to measure the impact of the information intervention in the medium term (three months after graduation). Below, I check if the difference in actions taken by the information and no-information groups after graduation is equivalent to the difference in intentions of the information group before and after the information intervention. This will test the usefulness of the data collected. I find that the results of the survey indeed translate into corresponding differences in sample actions.

4.1.1 Expectation Elicitation

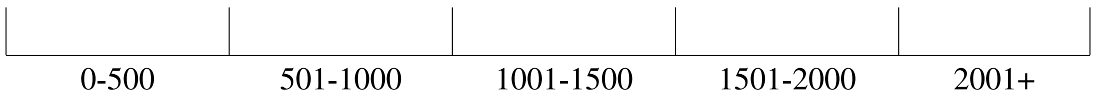
There are a variety of ways in which one can elicit expectations from an individual about future events. Manski (2004) provides a useful summary of the subjective expectations literature and methods while Delavande et al. (2011a,b). show the relative advantages of different elicitation procedures. After extensive pilot testing, it was found that the elicitation method most natural for this sample was a "static P.D.F".

The static P.D.F. elicitation method asks individuals to provide the probability density function that they believe holds for a particular variable. The elicitation method is static since the P.D.F. uses a number of discrete bins that do not change in size. Other elicitation methods such as a "dynamic P.D.F." first ask the individual for the support of their beliefs, by eliciting the maximum and minimum value of the variable of interest, then creates equally sized bins in response. The

[19]These students serve as a quasi-random comparison group as opposed to an explicit randomized control group. The data from the group of students who received the information intervention was collected a few of weeks before the data on those who did not receive the information. The classes were chosen based on availability at the time of the school visit and was not explicitly randomized. The visits to survey the comparison groups surveyed some of the classes in the same subjects in those schools that were not surveyed the first time. Appendix tables show that these students are statistically equivalent in terms of demographics, beliefs and occupational intentions.

static *P.D.F.* in this case affords at least two advantages: (1) it unifies the way that the elicitation procedure is explained to the sample, and (2) it allows for the same P.D.F. design to be used when providing information. This makes comparisons between the individual's P.D.F. and the income distribution P.D.F. provided to them much easier to do.

In all of the elicitations, the individual was asked to consider five bins of the following form:



Each bin covered a certain range of monthly income in Egyptian pounds[20], ranging from zero to five hundred pounds and increasing by five hundred pounds per bin until the last bin which was open-ended. The students were told that they had a total of 100 points to distribute across the five bins. They were to allocate more points to the bins that they felt were more likely to reflect the outcome for them if they successfully pursued a particular occupation after they graduated. For instance, if they were quite sure that if they pursued wage work that they would receive a wage between 600 and 800 pounds then they could allocate all one hundred points to that particular bin and leave the rest with 0 points. If on the other hand they were unsure about their expected income and saw that each bin was just as likely they could allocate 20 points to each bin. They were also able to freely allocate their points in any way between these two extremes.

This elicitation technique was explained to them using an example about the expected temperature on a given day next week. A similar set of five bins were introduced with different temperature ranges being provided for each bin and points were allocated across the bins in different ways to explain the concept. Nonetheless some students had trouble with this set up. About 20% of the students were unable to correctly allocate exactly 100 points across all the bins. Given the educational level of these students this was not unexpected. Responses were normalized for any allocation that did not add up to 100 in order to be comparable to the rest of the sample.

4.1.2 Information Intervention

Information can be a powerful tool in changing people's behavior. Problematically, it is sometimes too easy to provide information in a way that is unclear or misleading. In cases in which information is misleading, it might prompt individuals to make decisions that are against their best interest. For this reason, it is especially important to find the clearest way to provide information

[20]The exchange rate at the time of the study was approximately 7 EGP for 1 USD. The bins reflected expected monthly income as that is the primary way people measure income in the local context (as opposed to the American custom of considering annual income).

so that individuals are able to use it optimally. Some criticisms of earlier work in information provision include that labor market returns were conveyed in a manner that may have ignored both the natural heterogeneity of returns as well as the myriad of estimation issues that can plague estimates of returns.

To that end there are two main novelties in how I provide information. First, instead of providing them with the "return" to a particular occupation or simply providing the mean and standard deviation of income in the occupation, I provide the entire distribution of income to students. This allows for the students to appreciate the natural heterogeneity in income found in different occupations. Secondly students were told a story that explained the concept of the endogeneity of income in the different occupations. This was done to make clear that although one occupation may seem to be more lucrative than another it is possible that this is due in part to the selection of more able individuals into that occupation.

The data were taken from the recent Egypt Labor Market Panel Survey (2012). This data set surveys over 50,000 individuals in Egypt during 2012. It is the first large scale survey conducted after the Egyptian Revolution of 2011. The survey provides information on a variety of variables including labor market status and earnings for a representative sample of Egyptians. The data provided to the students restricts the sample to those that graduated from 3-year vocational schools and were between the ages of 18 and 30. Unfortunately, there was not enough data to split the sample down further by either geography or gender. The students were informed of these limitations to the data provided.

The occupations were split into Wage-Work, Self-Employment, and Business Owners (having at least 1 employee). The students were provided the average income for each occupation as well as the distribution of income in that occupation. Figure 1 below shows how the distribution of incomes was presented. In addition to information on income, students were also told what proportion of the working population in this sample populated each of the three occupations. Unfortunately, unemployment statistics for each occupation were not separately available and so overall labor force participation rates were provided to the students (separately for each gender).

Figure 1 shows that the distribution of income for wage workers leans left with 45% of individuals earning between 500 and 1000 Egyptian pounds a month. A large proportion manages to do better with 20% earning between 1001 and 1500LE, 10% earning between 1501 and 2000LE and 7% earning more than 2000LE. Appendix A shows the other two distributions for those were are self-employed. After the distributions were presented to the students the main points differentiating the occupations were discussed. This included the greater variability of earnings in self-employment as well as the greater mass of individuals in self-employment occupying the lowest bin when compared to any of the other graphs. Appendix B has the full transcript of the information intervention.

4.2 Local Context and Sample Demographics

Al-Fayoum is one of 27 governorates in Egypt. It ranks fourteenth in population with approximately 2.9 million residents. Despite being about 50 miles from Cairo it is known as one of the poorest areas in Egypt. The Egyptian Revolution of 2011 has had a significant impact on the national economy with the economy of Al-Fayoum being no exception to the general decline. The revolution has also introduced a fair bit of uncertainty to the marketplace. This has been referenced in numerous news articles and is anecdotally clear to residents and students of the schools. Egypt has been marred by high unemployment, especially high youth unemployment. Official statistics claim unemployment to be around 25%, while unofficial sources claim numbers that are much higher.

Table 1 showcases the basic demographic information collected from the students. It shows a number of interesting results. First, I find that the great majority (93%) of males are single while only 40% of women are single. This speaks to the large female labor force participation problem that is prevalent in Egypt. Next, I find that the students generally come from large families (an average of 4.17 siblings), and many have parents are who uneducated. One third of the students' fathers are illiterate, while more than half of mothers are illiterate. Only a small proportion of students have parents who are college educated.

Many students come from families in which the fathers are unemployed, which is to be expected in these trying times in Egypt, however, the extent to which this is true is somewhat shocking. Overall, Table 1 tells us that the sample generally comes from the lower end of the demographic pyramid in Egypt.

About a quarter of the sample is labeled as highly risk averse. These distinction is made using the students' responses to two questions about risk appetite. The first question asks the student to rate their willingness to take risks on a scale from 1-10. The second question gives them the option of 6 different lotteries and asks them which of the six they would prefer. Those students that responded consistently risk averse across the two questions were labeled as highly risk averse.

Appendix Table 10 shows that the basic demographics are equivalent for the comparison group within each gender.

5 Results

The survey and information experiment provides new and interesting data that allow me to derive new estimates of the determinants of occupational choice. This section will review the results from these data. I begin by looking at the baseline beliefs elicited from the sample in section 5.1. In section 5.2, I show how the information and credit interventions leads to aggregate changes in occupational intentions. I show that these aggregate shifts cannot be easily explained by looking at changes in aggregate beliefs. In section 5.3, I begin to estimate the determinants of occupational choice by first using the standard (yet potentially biased) techniques in the literate for comparison; in section 5.4, I outline how the information treatment changes expectations in a logical manner. Section 5.5 presents my unbiased results on the determinants of occupational choice and re-examines the aggregate impacts of the interventions in light of my new estimates. In section 5.6, I discuss the impact of the interventions on sector allocation, while section 5.7 demonstrates how information affects the overall efficacy of the credit intervention. Section 5.8 considers the welfare effects of each intervention, and section 5.9 estimates the medium term effects of the information intervention using the comparison group.

5.1 Baseline Beliefs

I elicit student's beliefs about a variety of observable expected outcomes in different occupations. These include: (1) income, (2) variance of income, (3) probability of successfully finding a job/starting a business, and (4) non-work income.[21] I elicit these expectations in wage work and self-employment and again in self-employment given the hypothetical availability of an enterprise loan. For the case in which individuals would leave the labor force, I only elicit their expected mean income.[22] The data is summarized in Table1.

The summary statistics provide some interesting insights into the beliefs of the sample. First, there are significant differences in average earning expectations between men and women. This occurs for both personal earnings and expected population earnings in both wage work and self-employment. When compared to actual population earnings found in the data (918 for Wage Work and 1024 in Self-Employment), men slightly over estimate population earnings while women severely underestimate them in wage work. Looking to expected personal earnings, similar patterns are found with men expecting higher earnings in general in each of the occupations. The availability of the loan leads to a higher expected income in self-employment for both men and women. Interestingly, the variability of income seems to be slightly higher in wage work. Income in inactivity is found to be quite low overall.

Expected probabilities of finding a job show that men only expect to be able to find a job within the next year with 49% probability while women only expect a 41% chance of finding a job. These seem quite pessimistic and speak to the large youth unemployment problem in Egypt[23]. Men and

[21] I also collect data about perceived relative and absolute skill as well as data about expected spousal quality in each occupation, as these are also potentially observable inputs that can impact occupational choice. I find no effects on these variables and so for clarity do not discuss them in detail.

[22] Although it may have been interesting to collect information about the variance of expected income for the case where they leave the labor force, pilot testing found this idea to be difficult for the students.

[23] It is possible that these low job finding probabilities are largely accurate. Unfortunately, official statistics about

women have similar beliefs about their ability to open a successful business with both genders believing they would have an average of a 64% success rate.

Overall, I find that baseline expectations of the sample are seemingly reasonable. There seems to be a slight underestimation of the population income in wage work and an overestimation of the population income in self-employment relative to that found in the representative survey.

Appendix 11 shows that baseline beliefs are equivalent for the comparison group within each gender.

Table 1 shows how those expectations were translated into baseline occupational intentions by the sample. About 43% of individuals intend to pursue wage work, 34% self-employment, and 23% plan to leave the labor force. Women in particular more than double the proportion of men who plan to leave the labor force.

Appendix 11 shows that baseline occupation intentions are equivalent for the comparison group within each gender.

5.2 Aggregate Updating of Occupational Intentions

It is interesting to consider how these baseline beliefs change in response to the credit and information interventions described above. Below, I first show how occupational intentions were impacted by the offer of an enterprise loan, then how they change after the students are subjected to the information intervention and finally how both credit and information impact occupational intentions as compared to baseline intentions.

5.2.1 Credit Constraints

Table 2 shows the impact of each of the interventions on the aggregate occupational intentions of the sample. Each cell of the table is the coefficient on a regression looking at how the intention to pursue each occupation changes (compared to baseline) when subjected to the particular intervention. Column 1 provides the baseline occupational intentions of the sample. Column 2 shows how these intentions change when the opportunity to take out a loan to start a new business is provided. The offer of credit causes significant shifts into self-employment by the sample. There is a nine percentage point increase in the average probability to pursue self-employment which equates to a 26.7% increase in the proportion of the sample pursuing self-employment. This influx of nine percentage points is coming from a decrease of 4.9 percentage points in the proportion pursuing wage work and 4.1 percentage points in the proportion intending to leave the labor force.

unemployment in Egypt are low quality. Follow up data on the students three months after graduation show that 88% of students are not working.

If accurate this would imply that the credit intervention would increase labor force participation by 5.3%.

5.2.2 Information Intervention

Next, I consider the effect the information intervention had on the sample as a whole. The model is unable to provide any predictions about aggregate effects of my particular information intervention and instead provides individual level predictions about how changes in expectations will result in changes in intentions. The results in column 2 show that at the aggregate level information induces a shift in the the proportion of individuals that want to pursue self-employment with an increase of 2.2%. This shift is coming primarily at the cost of a decrease in the proportion of individuals who were originally intending to leave the labor force after they graduate.

Similar to providing credit, these results show that providing information can also shift individual's occupation intentions. Yet it would be incorrect to say that any type of information intervention is sufficient to do so. The ability to affect change in occupational intentions with information depends on exactly how that information shifts *expectations*.

5.2.3 Information & Credit

Column 4 shows how occupational intentions change relative to baseline after the individuals were provided with the information intervention and were also offered the opportunity for a loan to start a business. It shows that information and credit lead to a large increase in the proportion who plan to pursue self-employment with a corresponding decreases in the proportion who intend to pursue wage work or leave the labor force.

Interestingly, credit and information together induce changes from baseline intentions that are different from what one would expect if they simply added up the effects of credit alone and the effects of information alone. As shown above, credit alone leads to a shift towards self-employment, while information alone leads to a similar yet smaller shift. While in the case where credit is available, information provision induces individuals to leave self-employment for wage work instead. Column 5 shows this most clearly by contrasting the intentions of the sample when they are provided credit before the information intervention with their intentions when they are provided credit after the information intervention (column 5 considers the changes from column 2 to column 4).

Column 5 shows that *given credit,* the information intervention leads to a shifting of individuals away from self-employment and into wage work. This result is quite interesting as it shows that the *same information* is leading to shifts in occupational intentions in opposite directions depending on whether or not credit is available.

5.2.4 Explaining the changes

The theoretical model provides a framework to understand why these changes are occurring. The model explains changes in occupational intentions as the culmination of changes in the expectations of outcomes in the occupations and how that impacts utilities in the occupations. But it's not clear which outcomes are most important. Is it the case that the mean of income is the most important determinant, or the probability of success, or any of many other aspects of a particular occupation? Knowing which are the main determinants of occupational choice can be quite important for a social planner who may be interested in shifting individuals into and out of an occupation. If it is the case that the mean of income is the most important determinant then this might lead to changes in the tax system while if the issue is in the volatility of income then helping to develop more robust insurance schemes may be ideal, etc.

Simply looking at the changes in aggregate expectations is not sufficiently informative to allow me to explain the changes in aggregate intentions. Let's take the impact of the information on occupational intentions as an example. Table 3 shows how the information impacts expectations at the aggregate level. In general, the table shows small and noisy impacts on expectations across the occupations. For instance, expected income seems to go down in all three occupations by similar amounts. The one exception is that the expected probability of success in wage work increases by 6.5%.

What this table illustrates is that it is unclear which of these changes are driving the aggregate impacts I find in table 2. For this reason it is important to estimate the determinants of occupational choice to better understand what types of things induce individuals to change their occupational intentions. The following subsections will estimate the determinants of occupational choice and then I will re-examine how the changes in expectations rationalize the changes in intentions found.

5.3 Earlier Modeling Strategies

Before delving directly into the estimates of the determinants of occupational choice using the information intervention, it is useful to consider earlier modeling strategies for comparison.[24] Many other papers have considered the determinants of occupational choice, but most make strong assumptions about the expectations that individuals hold about future outcomes. Generally, some form of "Rational Expectations" is assumed in which the expectations that individuals have are derived using some sort of functional form that takes into account what the econometrician would calculate their expectation to be. This assumption is restrictive and can be relaxed by collecting the individual subjective expectations that are collected in this paper. I can directly use the expectations elicited by the sample to understand how their expectations help determine their occupational

[24]For a detailed theoretical break down of the implications of the different strategies see Wiswall and Zafar (2013)

intentions.

Table 4 shows the results of a simple regression that would serve as the starting point for most studies of this sort. It simply regresses the elicited intention to pursue an occupation, in this case wage work, on the expected outcomes in that occupation, like income, the probability of finding a job, and the variance of income. It also interacts a binary for highly risk averse individuals to allow for them to differ in their weighting of the riskiness of an occupation. I take logs of both the dependent and independent variables to fit the model, assist with interpretation of the coefficients, and to keep them easily comparable to the other regressions throughout this paper.[25]

Table 4 shows that expected income in an occupation is positively correlated with the intention to pursue that occupation. In particular, it shows that a 10% increase in expected income in wage work is associated with a 0.68% increase in the intention to pursue wage work. It also shows that the expected probability of success is very strongly associated with the intention to pursue an occupation, especially for those who are risk averse. It shows no effects on the expected variance in wage work.

Although the results of this regression seem to make sense, it is biased in a number of ways. First, it does not consider the expectations of outcomes in other occupations. This is an important omission as the choice of occupation is not done in a vacuum, but instead, by contrasting expected outcomes in one occupation with expected outcomes in the other occupations. The regression is missing the intention to pursue self-employment, an omitted variable that is negatively correlated with the dependent variable as well as negatively correlated with many of the independent variables leading to a positive omitted variable bias.

Table 5 attempts to improve on these issues by considering the determinants of occupational choice using a first-difference strategy. Instead of ignoring the expected outcomes in all other occupations, it regresses the difference in intentions to pursue wage work relative to self-employment on the corresponding relative differences in expected income, success, and variance of income. This is the empirical counterpart to equation 2 in the model.

Table 5's results are quite different from those found in Table 4. The coefficient on the effect of expected income is halved and no longer significant. Similarly, the coefficient on the effect of the probability of success changes from being strongly positive to being negative yet insignificant. Finally, there is now an effect from the variance of income where as the relative variance in wage work vs. self employment increases by 10%, the intention to pursue wage work relative to self-employment decreases by 0.36%.

Comparing Tables 4 & 5, it becomes clear that the simple regression was biased by not accounting for the expected outcomes in other occupations. Nonetheless, as the theoretical section

[25]Taking logs is problematic in cases where there are many zeros. For this reason, all probabilities were changes from 0 to 0.001 and from 1 to 0.999 as is the standard in these kinds of models (Blass et al., 2010).

of the paper outlines, Table 5 also suffers from omitted variable bias. In particular, it is still not controlling for unobservables, like tastes for non-pecuniary aspects of an occupation. These unobservables can have important effects on the regression. For instance, let us consider the taste for being one's own boss. This taste will clearly be positively correlated with the intention to pursue self-employment. By construct, that means that it will be negatively correlated with the dependent variable in table 5. At the same time, it is not unreasonable to imagine that taste for being one's own boss is correlated with one's expectations about income in self-employment. It is possible that it is positively correlated (those that want to be their own bosses believe that they'd be better at business in general) or negatively correlated (those that want to be their own bosses are more willing to take self-employment jobs even if the income in self-employment is low). In this case, not only is it reasonable to expect that unobserved taste can bias the results of the regression, it is not even clear in which direction that bias will be.

For this reason, it is important to find a way to control for these unobservables. As the theoretical section outlines one way to do so is to shock an individual into changing their occupational expectations and intentions. This allows me to use the panel of beliefs to difference out these time-invariant taste parameters. Doing so takes care of the omitted variable bias that is plaguing the first two regressions and provides unbiased results of the determinants of occupational choice. Before proceeding to the unbiased regression, I outline how the information intervention managed to change the expectations of the sample.

5.4 Effects of the Information Intervention on Beliefs

In this section, I show how the information treatment affected individual subjective expectations. Much has been written about how people form and update their expectations including Delavande (2008b); Nyarko and Schotter (2002); Stinebrickner and Stinebrickner (2012); Wiswall and Zafar (2011); Zafar (2011b). In this paper, it is not necessary to consider how the expectations are formed exactly nor the process in which they are updated. What is necessary is simply for some expectations to be updated. Measuring the changes in expectations and corresponding changes in intentions to pursue an occupation is sufficient to estimate the determinants of occupational choice.

Figure 4 shows how the information intervention impacts the sample's beliefs about the average income for the population in wage work. The blue line shows the true value of the population average (as found in the population survey) subtracted from the sample's initial beliefs about the population average. It shows that overall, the sample underestimates the population average in wage work. The red line shows the true value subtracted from the sample's beliefs after the information intervention. It shows a clear shift towards zero, the difference in distributions is statistically significant at less than a 1% level. This proves that the information treatment successfully manages to impart information about the population average in the data to some of the sample.

Similarly, Figure 5 shows that many people incorrectly estimates the average population earnings in self-employment in both directions. The red line once again shows how after the information intervention, individuals revise their population beliefs leading to a much larger mass around the "zero error" area. These two distributions are also statistically different at less than the 1% level. Overall, this shows that the information intervention successfully imparts information about population averages that the sample found credible.

5.4.1 How Shifts in Population Beliefs Induce Shifts in Personal Expectations

Given the shifts in the sample's beliefs about population averages, it is now worth exploring if those shifts result in changes to personal expectations. Table 6 does that by regressing the log revision in the population average on the log revision in expected personal income. I find that a 10% increase in the population average in wage work leads to a 1.88% increase in the average expected personal income. Similarly, a 10% increase in the population average in self-employment leads to a 2.0% increase in the average expected personal income in self-employment.

These regressions show that individuals logically revise their beliefs about personal outcomes when faced with information about general outcomes. Nonetheless, these estimates show that revisions aren't perfectly elastic implying that much more goes into how individuals determine their expected self earnings in addition to the type of information I provide.

5.5 The Determinants of Occupational Choice

Section 5.4 showed that the information provided to the students shifts their beliefs about population averages. Those shifts in beliefs resulted in changes to expectations about personal outcomes in the various occupations. Following Wiswall and Zafar (2013), and as explained in section 3.4, I can use these changes in expectations to improve my earlier attempts at estimating the determinants of occupational choice. By using a double difference strategy, I can remove all of the time invariant unobservable parameters that were biasing the earlier results. The exact expression for the corrected regression is:

$$(lnY_{dt'} - lnY_{d't'}) - (lnY_{dt} - lnY_{d't}) =$$

$$= \beta_1[(E[u(X_d|\Omega_{t'})] - E[u(X_{d'}|\Omega_{t'})]) - (E[u(X_d|\Omega_t)] - E[u(X_{d'}|\Omega_t)])] \tag{6}$$

In words, this expression regresses the change in the log intention gap between two occupations before and after the information intervention on the change in the expectations gap between those occupations. By using my panel of expectations, before and after the information intervention I am able to difference out all time-invariant characteristics, both observables and unobservables.

The assumption that taste parameters are time invariant is not unreasonable in the current context. Unlike most other data sets, the difference in time between the two observations from the same individual is quite short - only a matter of hours. Within these hours, the only thing that transpires is that the individual is subjected to the information treatment outlined above. That information treatment provides information only on the type of observable characteristics of occupations that are collected in the survey. This leads me to expect that any changes induced by the information treatment will be changes in the observable outcomes that I am collecting[26] and that the unobserved taste parameters would in fact be time-invariant.

Table 7 shows the results of the unbiased regression of the determinants of occupational choice. Column 1 shows that as the relative gap in expected earnings between wages and self-employment increases by 10%, then the individual intention to pursue wage work relative to self-employment increases by 0.45%. This supports the prediction of the model, yet the magnitude is surprisingly small, leading to the conclusion that the relationship between wages and occupational choice is quite inelastic. Much more must enter the occupational decisions of the individual students.

Columns 2 and 3 consider other important determinants like the variance of income and probability of success in the occupations. All coefficients in column 2 go in the direction expected in the model, but are small, and the coefficients on variance and the probability of success are not statistically significant. On the other hand, when I interact these two variables with a dummy for those who are highly risk averse, I find that, as predicted, risk averse individuals are much more sensitive to changes in the expectations of the risky aspects of an occupation.

When compared to the results from Tables 4 & 5, these results are a bit more intuitive. They support all of the predictions of the theoretical model and have interesting implications. The results suggest that different types of people will focus in on different aspects of new information.

5.5.1 Re-Examining the Aggregate Impacts

Now that Table 7 provides accurate estimates of the determinants of occupational choice, it is worthwhile to re-examine the aggregate impacts of the different interventions that were found in Table 2. As the model outlined, the change in occupational intention is a function of the changes in many different potential outcomes in the various occupations. When estimating the aggregate changes in expectations ,the results will be much less precisely estimated than the changes in

[26]In addition to the observable outcomes of income, variance of income, and probability of success I also collect data on expected spousal quality, and perceived personal ability. It may be possible that data on the distribution of income leads to changes in these expectations as well which may impact occupational choice. Empirical analysis shows that these issues do not impact occupational choice and thus are omitted for clarity.

any one individual outcome. For this reason, although the changes in expectations can help us understand why we see the changes in intentions that we see, it will be rare that any of the changes in expectations will be statistically significant. Overall Table 7 summarizes how each change will impact intentions. Inducing a favorable change in expectations in one occupation relative to another will lead to a favorable change in intentions in that occupation relative to the other.

Column 2 of Table 2 shows a large shift of individual's towards self-employment after the availability of credit. The changes in aggregate expectations are clearly in favor of self-employment overall, as the model showed, the availability of credit weakly increases utility in self-employment alone. Table 12 shows how expectations in self-employment relative to wage work changed after the credit intervention. Particularly noteworthy is how the variance of income in self-employment increases when credit is available for the sample but decreases for those who are highly risk averse. This may speak to how highly risk averse individuals plan to utilize credit in their businesses.

Column 3 of Table 2 shows a shift away from inactivity and into employment. Table 13 shows how expectations of outcomes in employment, relative to inactivity, change due to the information intervention. This table, when compared to Table 3, provides a much clearer picture of why I see the aggregate impacts of information. Whereas Table 3 simply shows how aggregate income expectations in each of the occupations decreased after the information intervention, Table 13 shows that the expected income in employment *relative* to inactivity increases. It also shows that the probability of being successful in employment increases relative to inactivity. Table 7 shows that both of these changes would lead to the increases in the intention to pursue employment relative to inactivity and explains why I find the aggregate effect I found.

To explain why the shift into employment is happening almost completely into self-employment, one can consider the results from Table 14. It shows that, in the no-credit case, very little is changing regarding the relative standing of wage work vs. self-employment. Although it shows that the probability of success in wage work in increasing, it is not increasing for those who are highly risk averse and Table 7 shows that only the highly risk averse were sensitive to these types of changes. To explain why the sample shifts into self-employment, it's useful to notice that the constants in these regressions are all skewing towards self-employment. For instance the -511 constant term for income means that income in self-employment is 511LE greater than income in wage work.

Finally, Column 4 of Table 2 shows how credit & information changes expectations relative to baseline. Column 5 shows that *given* credit, information leads to a shift *away* from self-employment. Table 14 explains this by showing how expectations in wage work relative to self-employment changes when credit was available due to the information intervention. It shows a clear improvement in wage work relative to self-employment, especially for those who are highly risk averse.

As explained above, the *same* information can lead to changes in occupational intentions in opposite directions depending on whether or not credit is available. It is purely an empirical phenomenon, as the information is shifting expectations in the two cases differently. This can be seen in Table 14. Although many of the differences are not statistically different from one another at least one is- the relative probability of success. This shows how the information is impacting expectations in the two cases differently which leads to changes in intentions in different directions. This speaks to how information provision can interact with different situations and lead to varying responses to that information.

5.6 Effects on the Distribution of Risk Preferences Across Sectors

When considering the large shifts of individuals across sectors, it is natural to think about which people in particular are the ones being shifted. In this subsection, I show that because highly risk averse individuals react to information about risks more strongly than others, the new information leads them to shift away from self-employment more than those who are not highly risk averse.

More specifically, Table 8 shows how shifts into (and out of) self-employment differ by risk type in each intervention relative to baseline. Column 1 shows that the credit intervention induces a large overall shift into self-employment (as shown earlier in Table 2) and that that shift was not very different for highly risk averse individuals. Credit alone is able to shift both types of individuals into self-employment in close to equal proportions. The coefficient on the interaction is small and insignificant, leading to the conclusion that the offer of a loan has minimal impact on the types of individuals shifting into self-employment.

Column 2 shows how the information intervention has heterogeneous effects on the shift into self-employment depending on risk type. It shows that while the information intervention shifts the sample into self-employment, those who are highly risk averse barely shift in at all. The information intervention leads to a shift of 2.3 percentage points for the sample into self-employment but highly risk averse individuals are shifting in 1.9 percentage points less. Interestingly, highly risk averse people are still shifting into employment at the same rate but are instead shifting into wage work. This shows that people are sorting themselves into employment type based on risk preferences.

Column 3 considers how credit and information shifts individuals into self-employment compared to baseline occupational intentions. It shows a large shift into self-employment overall with a significantly smaller shift by the highly risk averse. Column 4 is potentially more informative as it shows how information shifts individuals *given credit.* It shows that the information intervention mitigated the impact of credit on shifting individuals into self-employment overall. Interestingly, it shows that even though the sample as a whole shifts away from self-employment those highly risk averse individuals shift out more than twice as strongly as the rest.

These results could be quite interesting in the case in which the allocation of risk averse individuals across sectors is important. In the next subsection, I outline why this is the case in occupational choice.

5.7 Information provision and the efficacy of policy tools

In cases in which constraints on credit or information bind, it can be expected that there will be frictions in the optimal allocation of individuals across sectors. For instance, if some people with high individual returns to capital are credit constrained, then they may be unable to pursue entrepreneurship. A social planner would want those with the highest return to capital to be entrepreneurs and so in this case may consider making credit more widely available. But to do so is costly and it is not clear that the credit, when made available, would go to those with the highest returns to capital. To fix this problem, the social planner would need to be able to both identify those who have high returns to capital and be able to provide it to them.

I posit that providing new information about the riskiness of different occupations leads to individuals self selecting into a more optimal usage of credit. This happens for two reasons (1) highly risk averse individuals are especially sensitive to changes in the expected riskiness of an occupation and the new information leads them to shift away from self-employment, and (2) highly risk averse individuals have lower average returns to capital. Together these two effects lead to the provision of new information improving the efficacy of the credit intervention with the average return to credit being higher in the information case relative to no information.

Proof of the first point is found in section 5.6 where it is shown that information leads to a strong shift away from self-employment for highly risk averse individuals. Evidence for the claim that risk averse individuals have lower average returns to capital can be found by examining how they report their expected changes in income when the loan becomes available.

It is natural to expect that loan availability leads to an increase in the expected income as well as the expected variance of income for those that will utilize the loan on their enterprise. This holds for the majority of the sample but for highly risk averse individuals we see a small increase in expected income, a small increase in the probability of success but a significant *decrease* in the expected variance of income (See Table 14). This implies that highly risk averse individuals would use the credit on safer, lower risk/lower reward, businesses. This would be less than optimal if the social planner is in search of "transformational entrepreneurs" to support.

More specifically, I can estimate how information impacts the expected returns to the credit intervention. To do so, I must first assume that the expectations that individuals have about outcomes in each occupation after the information intervention is correct on average[27]. Next, I use

[27] I explain this assumption further in the welfare section below.

these expectations to calculate the expected income with credit for those that would utilize it before the information treatment and contrast it with the expected income with credit for those that would utilize it after.

I find that the proportion of people interested in utilizing the credit intervention goes down after the information treatment from 30% to 24%. Importantly, the selection of those that do utilize it improves. In particular, the average expected income in self employment with credit before the information treatment is 11% lower than the average income after the intervention.

This shows that the information provided in this context is able to shift those who have lower returns to credit out of utilizing it while shifting those who have higher returns to credit into utilizing it. More simply, it shows that in certain cases information provision can lead to an improvement in the selection of individuals who utilize a policy intervention. This improved selection can have non-trivial impacts on the outcomes of that intervention improving the eventual outcomes by large amounts.

5.8 Welfare

It may be interesting to consider the welfare implications each of these interventions. By making a few assumptions I can calculate welfare before and after each intervention and consider how it changes.

To do so, I need to assume (1) a form of the utility function, and (2) that the expected outcomes in each of the occupations *after* the information intervention are the *true* outcomes that can be expected for these individuals. This second assumption is quite strong but can be understood somewhat as a "reverse rational expectations assumption". Rational Expectations normally assumes that the outcomes found in the data allow for the calculation of the average expectations that individuals hold. Whereas, in this case, I am assuming that the expectations that individuals hold will average out into the eventual outcomes of those individuals.

To calculate welfare I first assume the simplest utility function of the following form:

$$U_i = \sum_d Y_d * p_d * I_d$$

that is, welfare is equal to the probability of pursuing an occupation multiplied by the probability of successfully finding a job in that occupation (this is equal to 1 in the inactivity case) multiplied by the expected income in that occupation. Using this simple welfare function I can now consider three important cases: (1) how does credit affect welfare, (2) how does information provision improve individual welfare when credit is not available and (3) how does information provision improve individual welfare when credit is available.

In the first case, I look at expected utility using baseline intentions in the credit case vs the no credit case. I find that credit alone is able to improve expected utility by 10%. This improvement comes from an increase in expected income from the utilization of the credit and from more optimal sector allocation by the individuals across occupations.

In the second case, I use outcomes when credit is not available and contrast expected utility in the information vs no information case. I find that information alone is able to improve welfare by 6.5%. This improvement in welfare comes completely from improvements in sector allocation. In the third case, I use outcomes when credit is available and contrast expected utility in the information vs no information case. I find that given credit, information is able to improve welfare by 1.8%. Overall, information and credit improved expected utility by 12.3%.

These results are quite striking when considering the cost of information provision. If indeed providing information is able to induce improvements in welfare of this size, then the returns to providing information would be quite high as the costs of providing it are quite low. For instance, a reasonable estimate for the information intervention could be as low as $0.20 a student, whereas the cost of the credit intervention would be closer to $100 a student.

5.9 Medium-Term Effects of Information

This section provides the medium term effects of the information treatment utilizing data from both the treatment and comparison group several months after the initial survey. Student's were initially surveyed in March/April 2013 and graduated in late June 2013. Students were called in September 2013[28] to inquire about their labor market actions, which occupation they chose to pursue, and how they have been fairing at it.

This section serves as an important robustness check for the work above. If one believes the responses of the students on the survey, they can interpret the reported changes in occupational intentions to represent changes in occupational choices when the student enters the market and pursues an occupation. One might naturally be concerned about the accuracy of the students' responses considering the non-standardness of the questions and the uniqueness of this interaction they have had with researchers. To check on the usefulness of the data collected in the surveys, I consider how the sample responded to the information intervention. Since none of the students were actually offered the loan, I am unable to check how the survey responses about the loan would line up with their actions if they had the opportunity to take the loan.

The second overthrow of the Egyptian government led to difficulty in data collection and induced extra uncertainty to the students. The data that were collected show a dismal picture of how the Egyptian economy seems to young graduates. A staggering 88% of the sample are jobless

[28] The follow up calls were initially planned to happen in July 2013. Unfortunately the second overthrow of the Egyptian government started in July 2013 and so data collection was delayed until field activities were safe.

3 months after graduation, and of those with a job, 85% had it before they graduated from high school. These results give us a glimpse of the post-revolution economy.

Table 10 shows that baseline demographics, beliefs, and intentions are equal across the group that received the information and those that did not. Unfortunately the second overthrow of the Egyptian government led to great difficulty in data collection. The response rate of the sample to the follow-up survey was only about 25%. On the other hand, Table 15 shows that the attrition is evenly distributed across the sample.

Recall that the result of the information intervention at the aggregate level was a decrease by two percentage points in the proportion of individuals leaving the labor force and a corresponding increase by 2.2% in the proportion pursuing self-employment. If this is accurate then the follow up data should show a similar increase in the pursuit of employment by those that received the information treatment as compared to those who did not receive the information.

Table 9 presents the results from the follow up data. It finds a positive and significant effect on self-employment from the information intervention. Those that received the intervention were twelve percentage points more likely to be pursuing self-employment than those that did not receive the information. Even though the effect size is larger than expected size on the shift into self-employment, the corresponding shift out of the labor force is also similar in size, showing that the shift is coming from those that were planning to leave the labor force, exactly as expected. Overall, the effects aren't very precisely estimated and the impact expected from the survey lies within the confidence interval of the estimate from the follow-up data.

This gives support to the claim that the changes in occupational intentions I find in the survey credibly predict future actions by the sample. From a policy stand point, it is interesting that a simple information presentation was sufficient to induce individuals to pursue employment more. Unfortunately there is not sufficient variation to detect changes in actual employment due to the high jobless rate in post-revolution Egypt.

6 Discussion and Conclusion

6.1 Policy Implications

Constraints generally lead to individuals deviating from the optimal course of action that they would have taken in an unconstrained world. These deviations lead to decreases in efficiency at both the individual and collective levels. Policy makers strive to relax these constraints when possible, so as to allow the market to allocate resources (including human capital) more efficiently. This paper considers two important constraints that lead to potential deviations- credit and information constraints, how they change occupational choices, and what impacts this has on individual and aggregate outcomes.

I find that both credit and information constraints have non-trivial impacts on the decisions individuals have about which occupations to pursue and that relaxing these constraints leads to improvements in welfare. The information intervention improves individual welfare on average by 6.5% through improved sector allocation alone. The relaxation of credit constraints increases income and improves sector allocation leading to a 10% increase in welfare while relaxing both constraints leads to a 12.3% increase.

Most interestingly, the interaction of credit and information constraints with the heterogeneity in responses by highly risk averse individuals lead to important lessons about how information provision can impact the outcomes of a planned credit intervention. The differential response to information leads highly risk averse individuals to shift towards "safer" occupations on average, and out of using credit to start a small enterprise. Because those who are risk averse report lower returns to credit in general (by favoring lower risk/lower return investments), the average expected income for those utilizing the credit intervention rises by 12% relative to the no information case.

Many governments and international organizations contend that self-employment can be a solution to the growing unemployment problem found in numerous countries. If indeed that is the case policy makers will be interested in the different ways in which they can induce individuals to pursue entrepreneurship. These results show that credit provision can be a very powerful tool in contexts where credit constraints bind and that when coupled with information about the local market can induce a more efficient allocation of individuals across the different sectors of the economy.

If one can extend the lessons from this context to others, it shows that information provision can be a powerful tool in leading to improved outcomes both for individuals as well as for the more efficient utilization of policy interventions. Although the implications for other credit interventions are straight forward the lessons can extend to a variety of different programs. For instance, in the case in which individuals are information constrained and choosing between potentially investing in vocational training, this type of information intervention may lead individuals to pursue the training only if their expected returns to it are high. When the availability of spots for instruction are limited and costly, this type of improved selection can leave the individuals better off and leads to the training being more effective on average.

6.2 Conclusion

This paper is the first to consider the effects of both credit and information constraints on occupational choice. I develop a model of occupational choice that explicitly accounts for the two constraints and derive predictions that I test using data from a unique survey and information intervention in seven vocational high school in Egypt. I find support for my model's predictions: (1) credit availability leads to an increase in the proportion of the sample pursuing self-employment, (2) information that leads to a revision of expectations in an occupation lead to corresponding revisions in the intention to pursue that occupation and (3) highly risk averse individuals react more strongly to information about risk. This differential reaction has important consequences on the allocation of individuals across sectors as risk averse individuals strongly shift towards "safer" occupations. Follow up data three months after graduation confirms that the changes reported in occupational intentions translate directly into corresponding changes in actual occupations pursued.

This paper fits at the intersection of the literature on information constraints and how subjective expectations affect decision making and tackles the issue of occupational choice, a subject not previously considered using this lens. I expand on earlier work about information constraints and show how different behavioral preferences (in particular risk aversion) can affect the way people respond to different information. Combining insights from the literature on credit constraints, I provide novel insights into the occupational choice decision process of individuals about to enter the labor market.

Future work aims to consider the ways in which new information is transformed into changes in expectations. Other avenues of potentially fruitful research can be considering how different types of information lead to differential updating and impacts on occupational decision making.

References

Angelucci, Manuela, Dean Karlan, and Jonathan Zinman, "Win some lose some? Evidence from a randomized microcredit program placement experiment by Compartamos Banco," *National Bureau of Economic Research*, 2013.

Arcidiacono, Peter, V Joseph Hotz, and Songman Kang, "Modeling college major choices using elicited measures of expectations and counterfactuals," *Journal of Econometrics*, 2012, *166* (1), 3–16.

_ , _ , Arnaud Maurel, and Teresa Romano, "Recovering Ex Ante Returns and Preferences for Occupations using Subjective Expectations Data," 2013.

Attanasio, O. and K. Kaufmann, "Educational choices, subjective expectations, and credit constraints," 2009.

Attanasio, Orazio, Britta Augsburg, Ralph De Haas, Emla Fitzsimons, and Heike Harmgart, "Group lending or individual lending? Evidence from a randomised field experiment in Mongolia," 2011.

Augsburg, Britta, Ralph De Haas, Heike Harmgart, and Costas Meghir, "Microfinance at the margin: experimental evidence from Bosnia and Herzegovina," *Working Paper*, 2012.

Banerjee, Abhijit, "Contracting constraints, credit markets and economic development," *mimeo*, 2001.

_ , Esther Duflo, Rachel Glennerster, and Cynthia Kinnan, "The miracle of microfinance? Evidence from a randomized evaluation," *Working Paper*, 2013.

Banerjee, Abhijit V and Andrew F Newman, "Occupational choice and the process of development," *Journal of Political Economy*, 1993, pp. 274–298.

Bettinger, Eric P, Bridget Terry Long, Philip Oreopoulos, and Lisa Sanbonmatsu, "The Role of Application Assistance and Information in College Decisions: Results from the H&R Block Fafsa Experiment*," *The Quarterly Journal of Economics*, 2012, *127* (3), 1205–1242.

Bianchi, Milo and Matteo Bobba, "Liquidity, Risk, and Occupational Choices," *The Review of Economic Studies*, 2013, *80* (2), 491–511.

Blass, Asher A, Saul Lach, and Charles F Manski, "Using Elicited Choice Probabilities to Estimate Random Utility Models: Preferences for Electricty Reliability," *International Economic Review*, 2010, *51* (2), 421–440.

Blattman, Christopher, Nathan Fiala, DIW Berlin, and Sebastian Martinez, "Generating Skilled Self-Employment in Developing Countries: Experimental Evidence from Uganda," *The Quarterly Journal of Economics*, 2013.

Blau, Peter M, John W Gustad, Richard Jessor, and Herbert S Parnes, "Occupational choice: A conceptual framework," *Indus. & Lab. Rel. Rev.*, 1955, *9*, 531.

Boskin, Michael J, "A conditional logit model of occupational choice," *The Journal of Political Economy*, 1974, *82* (2), 389–398.

Carree, Martin A and A Roy Thurik, "The impact of entrepreneurship on economic growth," in "Handbook of Entrepreneurship Research," Springer, 2005, pp. 437–471.

Crépon, Bruno, Florencia Devoto, Esther Duflo, and William Parienté, "Impact of microcredit in rural areas of Morocco: Evidence from a Randomized Evaluation," *Working Paper*, 2011.

Delavande, A., "Pill, Patch, or Shot? Subjective Expectations and Brith Control Choice," *International Economic Review*, 2008, *49* (3), 999–1042.

Delavande, Adeline, "Measuring revisions to subjective expectations," *Journal of Risk and Uncertainty*, 2008, *36* (1), 43–82.

_ and Hans-Peter Kohler, "Subjective Expectations in the Context of HIV/AIDS in Malawi," *Demographic Research*, 2009, *20*, 817.

_ , Xavier Giné, and David McKenzie, "Eliciting probabilistic expectations with visual aids in developing countries: how sensitive are answers to variations in elicitation design?," *Journal of Applied Econometrics*, 2011, *26* (3), 479–497.

_ , _ , and _ , "Measuring subjective expectations in developing countries: A critical review and new evidence," *Journal of Development Economics*, 2011, *94* (2), 151–163.

Doepke, Matthias and Fabrizio Zilibotti, "Occupational choice and the spirit of capitalism," *The Quarterly Journal of Economics*, 2008, *123* (2), 747–793.

Dominitz, Jeff and Charles F Manski, "Using expectations data to study subjective income expectations," *Journal of the American Statistical Association*, 1997, *92* (439), 855–867.

Duflo, Esther and Emmanuel Saez, "The role of information and social interactions in retirement plan decisions: Evidence from a randomized experiment," *The Quarterly Journal of Economics*, 2003, *118* (3), 815–842.

Dupas, Pascaline, "Do teenagers respond to HIV risk information? Evidence from a field experiment in Kenya," *AEJ: Applied Economics*, 2011, *3* (1), 1–36.

Eeckhout, Jan and Boyan Jovanovic, "Occupational choice and development," *Journal of Economic Theory*, 2012, *147* (2), 657–683.

El-Gamal, Mahmoud, Mohamed El-Komi, Dean Karlan, and Adam Osman, "Bank-Insured RoSCA for Microfinance: Experimental Evidence in Poor Egyptian Villages," *Working Paper*, 2013.

Evans, David S and Boyan Jovanovic, "An estimated model of entrepreneurial choice under liquidity constraints," *The Journal of Political Economy*, 1989, pp. 808–827.

Ghatak, Maitreesh and Neville Nien-Huei Jiang, "A simple model of inequality, occupational choice, and development," *Journal of Development Economics*, 2002, *69* (1), 205–226.

Ginzberg, Eli, SW Ginsburg, S Axelrad, and JL Herma, "Occupational choice," *New York*, 1951.

Heckman, James J, Lance Lochner, and Christopher Taber, "Explaining rising wage inequality: Explorations with a dynamic general equilibrium model of labor earnings with heterogeneous agents," *Review of economic dynamics*, 1998, *1* (1), 1–58.

Jensen, R., "The (perceived) returns to education and the demand for schooling," *The Quarterly Journal of Economics*, 2010, *125* (2), 515.

Jeong, Hyeok and Robert M Townsend, "Sources of TFP growth: occupational choice and financial deepening," *Economic Theory*, 2007, *32* (1), 179–221.

Karlan, Dean and Jonathan Zinman, "Expanding credit access: Using randomized supply decisions to estimate the impacts," *Review of Financial Studies*, 2010, *23* (1), 433–464.

_ and _ , "Microcredit in theory and practice: using randomized credit scoring for impact evaluation," *Science*, 2011, *332* (6035), 1278–1284.

Kaufmann, K.M., "Understanding the Income Gradient in College Attendance in Mexico: The Role of Heterogeneity in Expected Returns to College," *Discussion Papers*, 2009.

Keats, Anthony, "Occupational choice in rural Kenya: Using subjective expectations data to measure credit and insurance constraints," *Working Paper*, 2012.

Ki-Moon, Ban, "Transforming Unemployed Youth Into Entrepreneurs Part of Solution to Global Crisis," June 2013. Remarks by Sectretary General Ban Ki-Moon at Thematic Debate on Development.

King, Allan G, "Occupational choice, risk aversion, and wealth," *Industrial and Labor Relations Review*, 1974, *27* (4), 586–596.

Klapper, Leora F and Inessa Love, "Entrepreneurship and development: the role of information asymmetries," *The World Bank Economic Review*, 2011, *25* (3), 448–455.

Malik, Adeel and Bassem Awadallah, "The economics of the Arab Spring," *World Development*, 2013.

Manski, C.F., "Measuring expectations," *Econometrica*, 2004, *72* (5), 1329–1376.

McFadden, Daniel and Kenneth Train, "Mixed MNL models for discrete response," *Journal of Applied Econometrics*, 2000, *15* (5), 447–470.

McKenzie, David and Christopher Woodruff, "Experimental evidence on returns to capital and access to finance in Mexico," *The World Bank Economic Review*, 2008, *22* (3), 457–482.

_ , John Gibson, and Steven Stillman, "A land of milk and honey with streets paved with gold: Do emigrants have over-optimistic expectations about incomes abroad?," *Journal of Development Economics*, 2013, *102*, 116–127.

Mel, Suresh De, David McKenzie, and Christopher Woodruff, "Getting credit to high return microentrepreneurs: The results of an information intervention," *The World Bank Economic Review*, 2011, *25* (3), 456–485.

_ , _ , and _ , "The demand for, and consequences of, formalization among informal firms in Sri Lanka," *Working Paper*, 2012.

Miller, Robert A, "Job matching and occupational choice," *The Journal of Political Economy*, 1984, pp. 1086–1120.

Naudé, Wim, "Entrepreneurship, developing countries, and development economics: new approaches and insights," *Small Business Economics*, 2010, *34* (1), 1–12.

Nguyen, T., "Information, role models and perceived returns to education: Experimental evidence from Madagascar," *Working Paper*, 2008.

Nyarko, Y. and A. Schotter, "An experimental study of belief learning using elicited beliefs," *Econometrica*, 2002, *70* (3), 971–1005.

Roy, Andrew Donald, "Some thoughts on the distribution of earnings," *Oxford Economic Papers*, 1951, *3* (2), 135–146.

Siow, Aloysius, "Occupational choice under uncertainty," *Econometrica: Journal of the Econometric Society*, 1984, pp. 631–645.

Staff, World Bank, *World Development Report 2013: Jobs*, Oxford University Press, Incorporated, 2013.

Stinebrickner, Todd and Ralph Stinebrickner, "Learning about academic ability and the college drop-out decision," *Journal of Labor Economics*, 2012.

_ **and** _ , "Academic performance and college dropout: Using longitudinal expectations data to estimate a learning model," *Journal of Labor Economics*, 2013.

Wiswall, M. and B. Zafar, "Belief Updating among College Students: Evidence from Experimental Variation in Information," *Working Paper*, 2011.

_ **and** _ , "Determinants of college major choice: identification using an information experiment," *Working Paper*, 2013.

Zafar, B., "Can Subjective Expectations Data be used in Choice Models? Evidence on Cognitive Biases," *Journal of Applied Econometrics*, 2011, *26* (3), 520–544.

_ , "How do College Students Form Expectations?," *Journal of Labor Economics*, 2011.

Tables

Table 1: Baseline Summary Statistics and Egyptian Population Comparison

	All	Male	Female	ELMPS
	(1)	(2)	(3)	(4)
Panel A				
Single	0.58	0.93	0.40	
Number of Siblings	4.17	4.13	4.19	
Father College Graduate	0.03	0.02	0.04	
Mother College Graduate	0.01	0.01	0.01	
Father Illiterate	0.33	0.40	0.30	
Mother Illiterate	0.55	0.59	0.52	
Father Working	0.61	0.53	0.65	
Mother Working	0.13	0.11	0.14	
Risk Averse	0.26	0.18	0.31	
Panel B				
Population Income in Wage Work	754	1076	584	918
Personal Income in Wage Work	963	1387	738	
Standard Deviation of Income in Wage Work	462	443	471	512
Probability of Finding a Job	0.44	0.49	0.41	
Population Income in Self-Employment	1449	1816	1254	1024
Personal Income in Self-Employment	1476	1791	1308	
Standard Deviation of Income in Self-Employment	460	442	470	668
Probability of Opening a Successful Business	0.64	0.64	0.64	
Income if Left Labor Force	191	208	181	
Unemployment Rate				0.20
Proportion out of Labor Force				0.50
Panel C				
Intention to Pursue Wage Work	.431	.491	.399	
Intention to Pursue Self-Employment	.337	.382	.312	
Intention to Leave Labor Force	.232	.127	.289	
N	1016	351	665	3168

Note: Risk Averse is a dummy representing those that answered consistently across questions that measure risk aversion: the first being a self report of willingness to take risks and the second being a choice between six lotteries. Panel B reports the sample's expectations of the variables at baseline. The standard deviation of income is calculated from the elicited PDF of expected income.Column 4 reports the means found in the Egypt Labor Market Panel Survey of 2012. The means are calculated after restricting the sample to those who have graduated from 3-year vocational high schools and are between the ages of 18 and 30.

Table 2: The Effects of Credit and Information on Occupational Intentions

Intent to Pursue:	Wage Work	Self Employment	Inactivity
	(1)	(2)	(3)
Panel A			
Baseline Intention (Constant)	0.431	0.337	0.231
	(0.005)	(0.005)	(0.005)
Impact of Credit	-0.049***	0.089***	-0.040***
	(0.009)	(0.009)	(0.008)
Impact of Information	-0.001	0.022**	-0.020**
	(0.008)	(0.007)	(0.007)
Impact of Credit and Information	-0.022**	0.058***	-0.036***
	(0.009)	(0.009)	(0.008)
Observations	4025	4025	4025
N	1016	1016	1016
Panel B			
Impact of Information *given* Credit	0.025**	-0.031***	-0.005
	(0.010)	(0.010)	(0.009)
Observations	1992	1992	1992
N	996	996	996

Note: The dependent variable is the response to the question "What is the probability that you will pursue occupation "X" when you graduate?". Responses that did not sum to 100% across all occupations were normalized to proportions of occupational intent. The bottom panel considers how information impacted occupational intentions *given* credit, essentially contrasting the second and fourth rows. Regressions include individual fixed effects and standard errors clustered at the individual level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 3: Changes in Aggregate Expectations After Information Intervention

	Income (1)	% Success (2)	Log Variance (3)
Wage Work	-66.1	0.065***	0.359
	(73.9)	(0.013)	(0.227)
N	1016	1016	992
Self Employment	-63.5	-0.012	0.292
	(128.8)	(0.012)	(0.204)
N	1016	1016	992
Inactivity	-81.3		
	(55.3)		
N	1016		

*Note:*Each cell reports the coefficient on a *post* dummy from a regression estimating the change in expectations of an outcome (column) in each occupation (row) after the information intervention. Regressions include individual fixed effects, with standard errors clustered at the individual level in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Table 4: Basic Regression on the Determinants of Occupational Choice

	Log Intention for Wage Work		
	(1)	(2)	(3)
Log Expected Income	0.069*	0.065	0.068*
	(0.034)	(0.038)	(0.035)
Log Expected Success in Finding Job		0.194**	0.153***
		(0.064)	(0.049)
Log Expected Success * Risk Aversion			0.182**
			(0.078)
Log Variance		-0.001	-0.005
		(0.007)	(0.009)
Log Variance * Risk Aversion			0.019
			(0.012)
N	1016	992	992

*Note:*Regressions consider baseline probability of pursuing wage work at baseline on baseline expectations of outcomes in wage work. Regressions include subject/class fixed effects, with standard errors clustered at the subject/class level in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Table 5: First Difference Regression on the Determinants of Occupational Choice

	Difference in Log Intention (Wage v Self Empl)		
	(1)	(2)	(3)
Difference in Log Income	0.028	0.034	0.031
	(0.058)	(0.061)	(0.061)
Difference in Log % Success		-0.000	-0.032
		(0.036)	(0.040)
Difference in Log % Success * Risk Aversion			0.138**
			(0.059)
Difference in Log Variance		-0.036**	-0.036**
		(0.013)	(0.012)
Difference in Log Variance * Risk Aversion			0.003
			(0.026)
N	1016	986	986

*Note:*Regressions consider the baseline probability of pursuing wage work *relative* to self-employment on baseline expectations of outcomes in wage work *relative* to self employment. Regressions include subject/class fixed effects, with standard errors clustered at the subject/class level in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Table 6: Revising Beliefs about Population and Self Earnings

	Log Revision in Expected Personal Income in:	
	Wage Work	Self-Employment
	(1)	(2)
Log Revision in Population Average	0.188***	0.200***
	(0.024)	(0.052)
N	1016	1016

*Note:*Regressions include individual fixed effects, with standard errors clustered at the individual level in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Table 7: Double Difference Regression on the Determinants of Occupational Choice

	Log Odds Revision (After-Before)		
	(1)	(2)	(3)
Log Relative Income Revision	0.045*	0.059*	0.060*
	(0.021)	(0.030)	(0.030)
Log % Success Revision		0.047	0.024
		(0.048)	(0.058)
Log % Success Revision * Risk Aversion			0.121**
			(0.055)
Log Variance Revision		-0.005	0.001
		(0.023)	(0.024)
Log Variance Revision * Risk Aversion			-0.033**
			(0.014)
N	1016	986	986

Note: Regressions consider the change, before and after the information intervention, of the probability of pursuing wage work *relative* to self-employment on the change in expectations of outcomes in wage work *relative* to self employment. Regressions include subject/class fixed effects, with standard errors clustered at the subject/class level in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Table 8: Changes in the Distribution of Risk Aversion Across Occupations

	Intention to Pursue Self-Employment			
	w/ Credit v. Baseline	w/Information v. Baseline	w/Info & Credit v. Baseline	w/Info & Credit v. Credit Only
	(1)	(2)	(3)	(4)
Effect of Intervention	0.089***	0.023***	0.064***	-0.020**
	(0.008)	(0.007)	(0.008)	(0.008)
Interacted with Risk Aversion	-0.014	-0.019*	-0.025**	-0.027**
	(0.014)	(0.012)	(0.013)	(0.013)
N	1016	1016	996	996

Note: Regressions in columns 1-3 consider how each intervention (column) impacted the reported probability of pursuing self-employment relative to the reported baseline probability. Column 4 instead considers the change relative to reported probability with credit available, essentially comparing columns 1 & 3. Regressions include individual fixed effects, with standard errors clustered at the individual level in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Table 9: Medium Term Effects of Information (Info vs No Info Groups)

	Pursing Self Employment		Left Labor Force	
	(1)	(2)	(3)	(4)
Information Treatment	0.115**	0.124**	-0.097^{+}	-0.092^{+}
	(0.055)	(0.053)	(0.061)	(0.063)
Treatment*Risk Aversion	0.047	0.023	0.014	0.006
	(0.108)	(0.107)	(0.112)	(0.115)
Controls	N	Y	N	Y
N	330	330	330	330

*Note:*Controls include gender, marital status, baseline income expectations, and parental education and work status. The sample includes 173 individuals who received the information treatment and 157 that did not. Standard errors clustered at the class level in parentheses, *** p<0.01, ** p<0.05, * p<0.1,$^{+}$p<0.15

Appendix Tables

Table 10: Demographics of Comparison Groups are Balanced

Dependent Variable:	In Information Group (=1)	
	Male	Female
	(1)	(2)
Father High School	0.041	-0.046
Mother High School	-0.082	0.056
Father Illiterate	-0.004	0.017
Mother Illiterate	-0.019	-0.010
Siblings	0.011	0.014*
Risk Averse	0.027	0.026
Father Working	-0.002	0.035
Mother Working	-0.001	0.073
Married	0.275*	0.064
F-Stat	0.91	1.48
P-Value	0.53	0.13
N	491	1096

*Note:*Risk Averse is a dummy representing those that answered consistently across questions that measure risk aversion: the first being a self report of willingness to take risks and the second being a choice between six lotteries.Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Table 11: Baseline Beliefs and Intentions of the Highly Risk Averse are Balanced

Dependent Variable:	Risk Averse(=1)		Risk Averse(=1)
	(1)		(2)
Population Income - Wage Work	-9.82e-08	Intention for Wage Work	-0.025
Population Income - Self-Employment	-2.24e-06	Intention for Self-Employment	-0.087
Personal Income - Wage Work	-1.04e-06		
Personal Income - Self-Employment	1.38e-06		
Personal Income - Inactivity	-3.37e-06		
F-Stat	0.26	F-Stat	0.55
P-Value	0.93	P-Value	0.57
N	1016	N	1016

*Note:*Risk Averse is a dummy representing those that answered consistently across questions that measure risk aversion: the first being a self report of willingness to take risks and the second being a choice between six lotteries.Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Table 12: Changes in Aggregate Expectations from Hypothetical Credit Intervention

	Wage Work - Self Employment:		
	Income	% Success	Variance
	(1)	(2)	(3)
Change from Baseline	-257	-0.03	-0.63**
	(254)	(0.07)	(0.26)
Change from Baseline * Risk Averse	-526	-0.03	0.91*
	(528)	(0.10)	(0.54)
Constant	-653	-0.47	0.08
	(111)	(0.03)	(0.11)
N	1016	1016	992

*Note:*Regressions consider changes in expectations of *relative* outcome (column) before and after the hypothetical credit intervention. Relative outcomes in this case are the expected outcome in wage work minus the the expected outcome in self-employment. Regressions include individual fixed effects, with standard errors clustered at the individual level in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Table 13: Changes in Aggregate Expectations from Information Intervention

	Employment - Inactivity		
	Income	% Success	Log Variance
	(1)	(2)	(3)
Change from Baseline	141	0.027**	0.187
	(256)	(0.011)	(0.214)
Change from Baseline * RA	-114	-0.002	-0.397
	(302)	(0.021)	(0.348)
Constant	932	0.541	10.84
	(96)	(0.005)	(0.085)
Obs	2032	2032	1984
N	1016	1016	992

Note: Regressions consider changes in expectations of *relative* outcome (column) before and after the information intervention. Relative outcomes in this case are the expected average outcome in employment minus the the expected outcomes in inactivity. Success in inactivity is assumed to equal 1 while variance is assume to equal 0. Regressions include individual fixed effects, with standard errors clustered at the individual level in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Table 14: Changes in Aggregate Expectations from Information Intervention

	Wage Work - Self Employment					
	No Credit			Credit		
	Income	% Success	Log Variance	Income	% Success	Log Variance
	(1)	(2)	(3)	(4)	(5)	(6)
Change from Baseline	1	0.103***	0.095	185	0.178***	0.239
	(164)	(0.018)	(0.274)	(174)	(0.021)	(0.310)
Change from Baseline * RA	-14	-0.097***	-0.239	115	-0.094**	0.034
	(325)	(0.036)	(0.310)	(340)	(0.045)	(0.499)
Constant	-511	-0.196	0.071	-511	-0.196	0.071
N	996	996	996	996	996	996

Note: Regressions consider changes in expectations of *relative* outcome (column) before and after the information intervention in cases with, and without, credit. Relative outcomes in this case are the expected outcome in wage work minus the the expected outcome in self-employment. Regressions include individual fixed effects, with standard errors clustered at the individual level in parentheses, *** p<0.01, ** p<0.05, * p<0.1

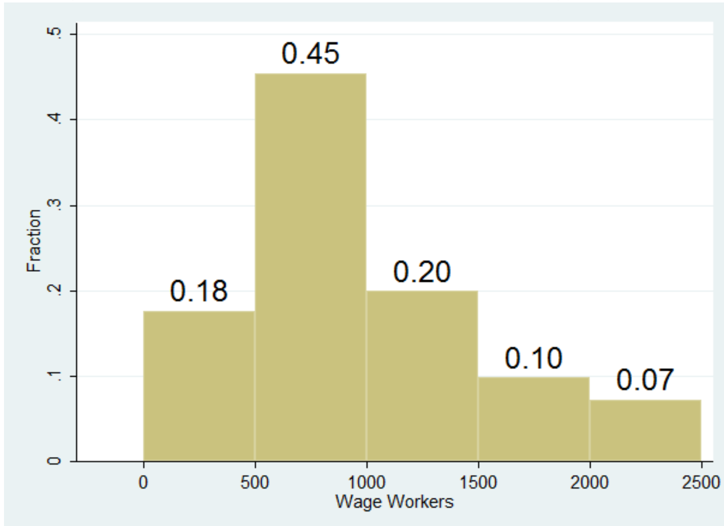
Table 15: Attrition on Follow-Up Survey

Dep. Var:	Surveyed(=1)	Dep. Var:	Surveyed(=1)	Dep. Var:	Surveyed(=1)
	(1)		(2)		(3)
Father High School	0.026	Population Income - Wage	-9.6e-06	Intention for Wage	0.017
Mother High School	0.033	Population Income - Self-Emp	-1.7e-05	Intention for Self-Emp	-0.045
Father Illiterate	-0.025	Personal Income - Wage	-7.08e-07		
Mother Illiterate	0.006	Personal Income - Self-Emp	3.95e-06		
Siblings	-0.007	Personal Income - Inactivity	-2.8e-06		
Father Working	-0.014				
Mother Working	-0.036				
Student Working	0.013				
Risk Averse	0.005				
F-Stat	0.93	F-Stat	0.86	F-Stat	0.54
P-Value	0.49	P-Value	0.51	P-Value	0.58
N	1574	N	1574	N	1574

*Note:*Risk Averse is a dummy representing those that answered consistently across questions that measure risk aversion: the first being a self report of willingness to take risks and the second being a choice between six lotteries. Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

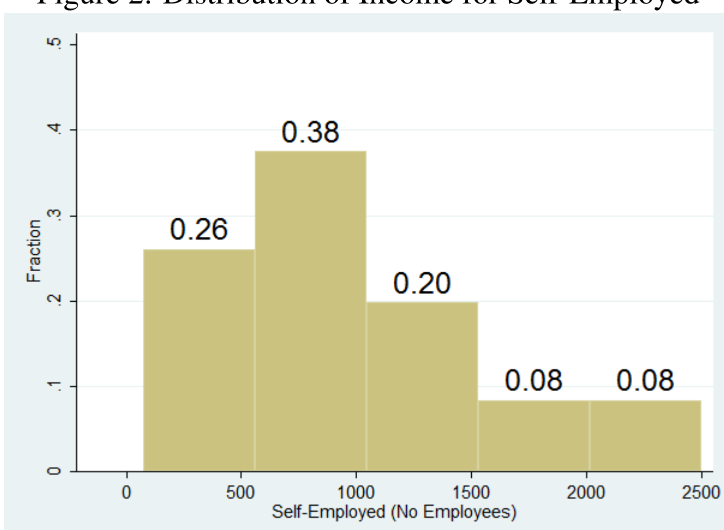
Figures

Figure 1: Distribution of Income for Wage Workers



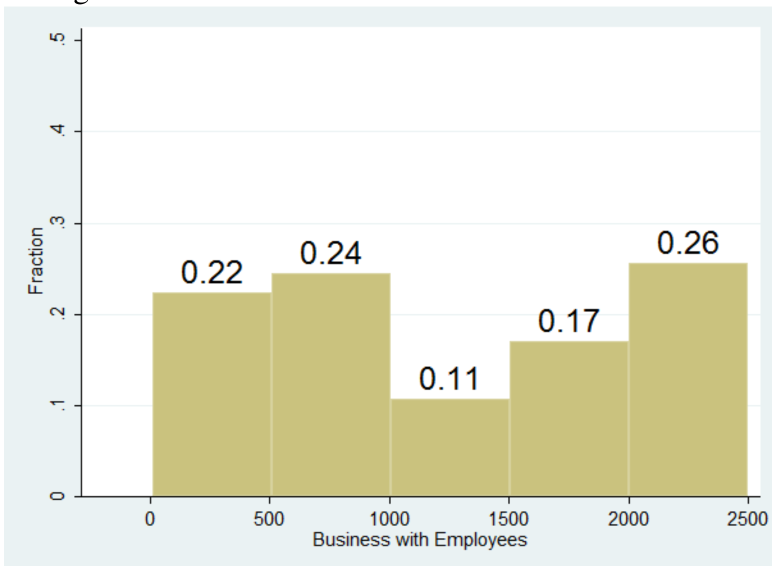
Average Income 918LE, 84.2% of earners

Figure 2: Distribution of Income for Self-Employed



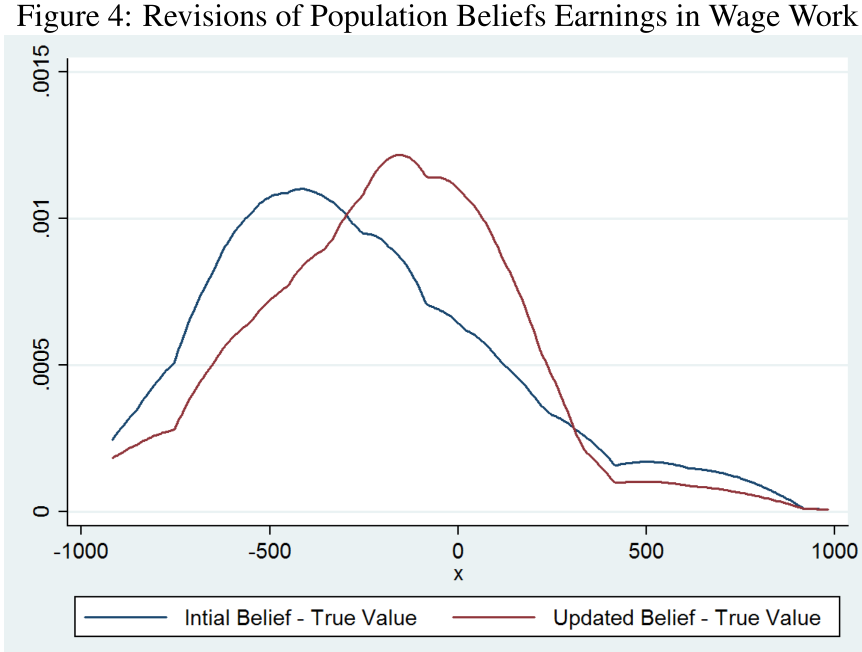
Average income 878 LE, 10.1% of earners

Figure 3: Distribution of Income for Business Owners



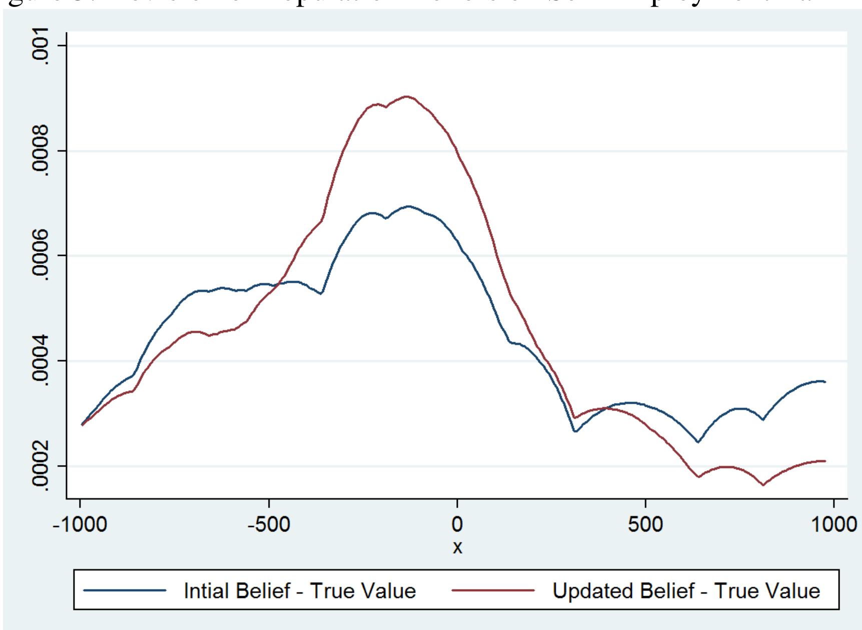
Average income 1318LE, 5.7% of earners

Figure 4: Revisions of Population Beliefs Earnings in Wage Work



Combined K-S P-Value: 0.000

Figure 5: Revision of Population Beliefs on Self Employment Earnings



Combined K-S P-Value: 0.000

Appendix - Information Script

We would like to provide you with a number of interesting findings from the most recent Egypt Labor Market Panel Survey, a country-wide survey done last year in Egypt. This survey allows us to see what the normal wages are in the labor market for different types of work without having to guess or just go by what we hear from a few people. This survey was conducted on nearly 50,000 people.

Here we are restricting our data to look at only those people who are similar to you. We are going to give you information about those that have graduated from 3-year vocational high schools. We are also only looking at people between the ages of 18-30 years old.

There are 3 different types of work that we want to speak about- Wage Employment, Self-Employment (Only employing oneself), Business Ownership (Employing others as well)

We will first give you the average wages for each of those three groups of people and then will be showing you the distribution of wages for each group as well as the proportion of the working population that can be found in each of the occupations. The average monthly income for wage-workers is 918LE The average monthly income for self-employed people is 878LE The average monthly income for business owners is 1125LE

Although these numbers are averages it's not the case that everyone make this much money a month. Some people in each of the groups will make more and some will make less. For this reason we will show you a graph of the proportion of people who make different incomes in each group on the next page.

Now we'd like to explain the three graphs you have in front of you. The first graph looks at the monthly incomes for those who are wage-workers. They make on average 918LE a month but some people make more and some make less. As you can see there are five columns and each column represents the proportion of people whose monthly income lies within the specified range. The first column shows that about 18% of people make between 0 and 500LE a month. The second columns shows that 45% of wage-workers make between 500 and 1000LE a month, the third column shows that 20% of people make between 1000 and 1500LE a month and the fourth column shows that 10% of people make between 1500 and 2000LE a month and the final column shows that 7% of people make more than 2000LE a month.

The second graph looks at the monthly incomes for those who are singularly self-employed. They make on average 878LE a month but some people make more and some make less. As you can see there are five columns and each column represents the proportion of people whose monthly income lies in within the specified range. The first column shows that about 26% of people make between 0 and 500LE a month. The second columns shows that 38% of wage-workers make between 500 and 1000LE a month, the third column shows that 20% of people make between 1000 and 1500LE a month and the fourth column shows that 8% of people make between 1500 and 2000LE a month and the final column shows that 8% of people make more than 2000LE a month.

The final graph looks at the monthly incomes for those who are employers. They make on average 1125LE a month but some people make more and some make less. As you can see there are five columns and each column represents the proportion of people whose monthly income lies in within the specified range. The first column shows that about 22% of people make between 0 and 500LE a month. The second columns shows that 24% of wage-workers make between 500 and 1000LE a month, the third column shows that 11% of people make between 1000 and 1500LE a month and the fourth column shows that 17% of people make between 1500 and 2000LE a month and the final column shows that 26% of people make more than 2000LE a month.

When we compare the three graphs we see that entrepreneurs seem to earn the most but at the same time this is being driven by a relatively small proportion of people who make lots of money. These are the people who were especially able to grow their business and see it succeed. The great majority of entrepreneurs earn similar amounts to those in the other occupations.

When we look for other differences between the three graphs we see that those who are self-employed are more likely to earn between 0 and 500LE than any other group. Similarly we see that there are more people in that low earnings bin from those who are entrepreneurs as opposed to wage workers. On the other hand those that are entrepreneurs have by far the greatest proportion of people who earn more than 2000LE, more than three times as many as those in wage work or singular self-employment. This implies that entrepreneurship in general has greater variability in earnings than wage work and hence potentially greater risk.

It's also important to note that the numbers we are seeing are the results of a choice made by people about the type of job that is best for them. As an example think about a very wage worker who makes 6,000LE a month and a very good entrepreneur that makes 8,000LE a month. It's not the case that if the entrepreneur decided to become a plumber that he would make 8,000LE or even that he would make 6,000LE a month. He might make even less because as he might not be a very good wage worker. Similarly a very good wage worker might not be able to be a good entrepreneur and would make less than 6,000LE a month if he tried. This goes to say that the data that we have are already the results of a choice made by people about what is best for them.

Finally, it's important to realize that not everyone who wants to work is able to do so. Our data tell us for men, out of every 100 people who graduated from vocational high school – 70% are working, 6% can't find work and 24% are not looking for work. For women 21% are working, 11% can't find work and 68% are not looking for work.